

Dissertation zur Erlangung des Doktorgrades der

Fakultät für Chemie und Pharmazie der

Ludwig-Maximilians-Universität München

Synthese und Eigenschaften modifizierter Nukleoside in DNA und RNA

Matthias Quirin Kurz

aus Fürstenfeldbruck

2020

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Herrn Professor Dr. Thomas Carell betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 21.07.2020

Matthias Kurz

(Matthias Kurz)

Dissertation eingereicht am: 28.05.2020

1. Gutachter: Prof. Dr. Thomas Carell

2. Gutachter: Dr. Stefanie Kellner

Mündliche Prüfung am: 17.06.2020

Publikationsliste

Teile dieser Arbeit wurden bereits publiziert und auf Konferenzbeiträgen vorgestellt.

Publikationen

- „*ALKBH5-induced demethylation of mono and dimethylated adenosine*“ Timm T. Ensfelder[#], Matthias Q. Kurz[#], Katharina Iwan[#], Simon Geiger, Sarah Matheisl, Markus Müller, Roland Beckmann and Thomas Carell, *ChemComm*, **2018**, 54, 8591-8593
- „*Non-canonical Bases in the Genome: The Regulatory Information Layer in DNA*“ Thomas Carell, Matthias Q. Kurz, Markus Müller, Martin Rossa, and Fabio Spada, *Angew. Chem. Int. Ed.* **2018**, 57, 4296– 4312

Die mit [#] markierten Autoren haben zu gleichen Teilen beigetragen.

Konferenzbeiträge

- „*Biological Significance and Occurrence of Modified Bases in RNA and DNA*“, SPP 1784 Meeting - Chemical Biology of Nativ Nucleic Acid Modifications (Dortmund, Juni 2018)
- „*A hypermodification in mRNA - The quest for function and sequence context of ms²i⁶A*“, 1st Symposium on Nucleic Acid Modifications (Mainz, September 2017)

Inhaltsverzeichnis

Zusammenfassung.....	IV
Summary.....	VII
1. Einleitung.....	1
1.1. Grundlagen des Aufbaus der DNA	1
1.2. Epigenetik.....	2
1.3. Die epigenetischen Basen in der DNA	2
1.3.1. 5-Methyl-2'-desoxycytidin.....	2
1.3.2. 5-Hydroxy-2'-desoxymethylcytidin	3
1.3.3. 5-Formyl-2'-desoxycytidin und 5-Carboxy-2'-desoxycytidin.....	3
1.3.4. Funktionen der epigenetischen Basen in der DNA.....	4
1.4. Sequenzierung der epigenetischen Basen in der DNA	5
1.4.1. Bisulfit Sequenzierung	6
1.4.1.1. BS-seq Methoden für hmdC	7
1.4.1.2. BS-seq Methoden für fdC.....	8
1.4.1.3. BS-seq Methoden für cadC.....	9
1.4.1.4. <i>Hairpin</i> Sequenzierung.....	10
1.4.2. Bisulfit freie Sequenzierung	11
1.4.2.1. Sequenzierung mit Hilfe von Endonukleasen	11
1.4.2.2. Chemische Derivatisierung und Anreicherung	11
1.4.2.3. Sequenzierung mit Hilfe von Antikörpern.....	12
1.5. Modifizierte Nukleoside in der RNA	13
1.5.1. Nukleosid-Modifikationen in mRNA	14
1.5.2. Das Epitranskriptom.....	14
1.5.2.1. <i>Writer</i> und <i>reader</i> von m ⁶ A.....	14
1.5.2.2. <i>Eraser</i> von m ⁶ A.....	15
1.5.3. Dealkylierung in tRNA	16
2. Aufgabenstellung	17
3. Antikörper gegen RNA- und DNA-Modifikationen	20
3.1. Allgemeines Hapten Design	20
3.2. Antikörper der erste Antikörper Generation	21
3.2.1. Synthese der Haptene PEG3-fdC und PEG3-Gm.....	21
3.2.2. EDC-Kopplung der Haptene 11 und 12 mit OVA.....	24
3.2.3. Charakterisierung der EDC Kopplungen der ersten Generation.....	25
3.2.4. Optimierung des ELISA-Screenings mit biotinylierten Oligonukleotiden	26
3.2.5. Festphasensynthese der Oligonukleotide für das ELISA- <i>Screening</i>	27
3.2.6. Optimierung der Dot Blot Bedingungen.....	30
3.2.6.1. Auswahl der Oligonukleotide	30
3.2.6.2. Auswahl der Crosslink-Methode	32
3.2.7. Auswertung der Dot Blots - Überstände erste Generation.....	32
3.3. Antikörper der zweiten Generation.....	35
3.3.1. Synthese des PEG1-fdC Haptens	35
3.3.2. EDC Kopplung der Haptene 11 und 17 mit OVA und cBSA.....	36

3.3.3. Charakterisierung der EDC Kopplungen zweiter Generation.....	37
3.3.4. Optimierung der Oligonukleotide für das ELISA-Screening	40
3.3.5. Auswertung der Dot Blots - Überstände zweiter Generation	41
3.4. Anwendungsmöglichkeiten der neuartigen Methode.....	43
4. Hairpin Oligonukleotide für die Sequenzierung von mdC, hmdC, fdC und cadC in dsDNA	45
4.1. Design der <i>Hairpin</i> Oligonukleotide	45
4.2. Synthese und Charakterisierung der <i>Hairpin</i> Oligonukleotide	48
4.3. Simultane Sequenzierung der epigenetischen Basen in dsDNA	50
4.3.1. Vorversuche zur Reaktivität von <i>SuperFly</i> mit fdC und fdU	52
4.3.2. Ausblick	53
5. Wasserlösliches IBX zur Oxidation von hmdC-Oligonukleotiden	54
5.1. Wasserlösliches IBX-Derivat mit negativer Ladung.....	54
5.1.1. Synthese des IBX-Derivates mit negativer Ladung (<i>p</i> -cIBX)	54
5.1.2. Oxidation von hmdC mittels <i>p</i> -cIBX im Oligonukleotid-Kontext	55
5.2. Wasserlösliches IBX-Derivat mit positiver Ladung.....	60
5.2.1. Synthese der IBX-Derivate mit positiver Ladung	61
5.2.1.1. Synthese des IBX-Amids.....	61
5.2.1.2. Synthese der Alkin-IBX Vorstufe.....	63
5.2.1.3. Synthese des SF ⁺ -IBX-Derivats.....	63
5.2.1.4. Synthese des PEG ⁺ -IBX-Derivats.....	64
5.2.2. Oxidation von hmdC mittels PEG ⁺ -IBX und IBX-Amid im Oligonukleotid-Kontext	66
6. Die Modifikation ms²A in RNA	70
6.1. Synthese eines d3-ms ² A	70
6.2. Anwendung des synthetisierten Standards	71
7. Veröffentlichte Arbeit.....	74
8. Experimenteller Teil	90
8.1. Chemische Synthesen	90
8.1.1. Allgemeine Arbeitsmethoden und Materialien.....	90
8.1.2. Durchführung der Synthesen.....	92
8.1.2.1. Synthese der Haptene	92
8.1.2.2. Synthese der wasserlöslichen IBX-Derivate.....	103
8.1.2.3. Synthese von d3-ms ² A.....	114
8.2. Oligonukleotidsynthese	117
8.2.1. Festphasensynthese	117
8.2.2. Entschützung und Aufreinigung.....	117
8.2.3. Charakterisierung	119
8.2.4. Oligonukleotid Sequenzen	119
8.3. Biochemie Methoden	121
8.3.1. Allgemeine Arbeitsmethoden	121
8.3.2. EDC-Kopplungen.....	122
8.3.2.1. Ovalbumin	122
8.3.2.2. kationisiertes BSA (cBSA)	123
8.3.3. SDS Polyacrylamid Gelelektrophorese (SDS-PAGE).....	124

Inhaltsverzeichnis

8.3.4. Dot Blots	125
8.3.5. Übersicht getesteter Hybridomzell-Überstände	125
8.3.6. Studien zur Oxidation eines hmdC-Oligonukleotid mit wasserlöslichen IBX-Derivaten	126
8.3.6.1. IBX Derivat mit negativer Ladung	126
8.3.6.2. IBX Derivate mit positiver Ladung (tertiäres Amin)	127
8.3.7. Reaktivitätsassay von SuperFly mit fdC- und fdU-Oligonukleotiden.	127
9. Abkürzungsverzeichnis	128
10.Literaturverzeichnis	132

Zusammenfassung

Multizelluläre Organismen teilen jeweils dieselben Sequenzinformationen, abgespeichert als DNA im Zellkern. Aus einem einzigen gemeinsamen Vorläufer heraus, den embryonalen Stammzellen, entwickeln sich hochspezialisierte Zelltypen. Diese Zelltypen unterscheiden sich in ihrer Beschaffenheit und nehmen unterschiedlichste Funktionen wahr. Das bedeutet, neben der Basensequenz muss es eine weitere Informationsebene geben, die von den unterschiedlichen Zelltypen individuell interpretiert werden muss und eine Kontrolle der Genexpression darstellt. Die DNA Modifikation mdC ist ein Beispiel für diese regulatorische Informationsebene. Mit den Cytosin-Modifikationen hmdC, fdC und caC wurden 2009 und 2011 drei weitere Nukleobasen der DNA entdeckt. Neben der Quantifizierung, stellt sich auch die Frage nach deren genauer Positionierung in genomischer DNA, um ihre wichtige biologische Rolle genauer zu verstehen. Dafür werden bis heute Wege zur Sequenzierung dieser epigenetischen Nukleobasen gesucht und bereits bestehende verbessert. Aber sowohl die Sequenzierung mittels Antikörper-Anreicherung, welche auch für RNA-Modifikationen sehr interessant ist, als auch die chemische Derivatisierung mit Natriumbisulfid (Bisulfit-Sequenzierung), bieten großes Optimierungspotential.

In der hier vorliegenden Dissertation, wurde eine neuartige Methode zur Herstellung von monoklonalen Antikörpern unter milden Reaktionsbedingungen etabliert. Mit dieser Methode sollte es nun möglich sein, neben empfindlichen Modifikationen an der Base, auch Antikörper gegen Modifikationen an der Ribose zu generieren. Es wurden dafür Haptene entwickelt, die mit Hilfe von EDC-Kopplung an Trägerproteine (OVA und cBSA) gekoppelt wurden. Zudem wurde das ELISA-*screening* optimiert, mittels der Benutzung von dafür synthetisierten modifizierten Oligonukleotiden. Mit der hier entwickelten Methode konnten erstmals monoklonale Antikörper gegen fdC hergestellt werden.

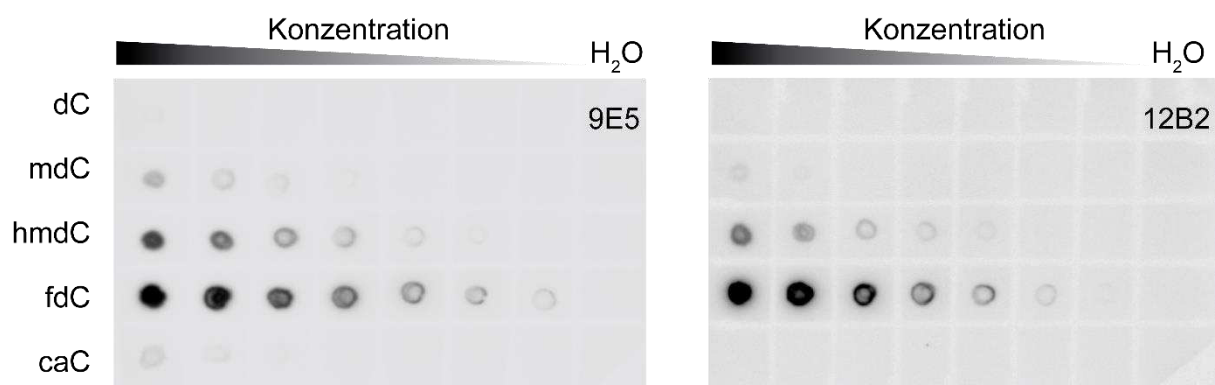


Abbildung I: Dot Blots von den in dieser Arbeit entwickelten monoklonalen Antikörper (9E5 und 12B2) gegen die Modifikation fdC. Die Antikörper wurden mit der neuartigen Methode und in Zusammenarbeit mit dem *Helmholtz Zentrum München* generiert. Dabei ist die Kreuzreaktivität der hier generierten fdC-Antikörper bezüglich dC und caC, im Vergleich zum kommerziellen Produkt deutlich geringer. (siehe Abbildung 39)

Diese sind mit der Spezifität des kommerziell erhältlichen polyklonalen fdC-Antikörpers der Firma *Active Motif* vergleichbar. Darüber hinaus ist die Kreuzreaktivität bezüglich dC und cadC, im Vergleich zum kommerziellen Produkt, deutlich geringer.

Ein weiterer Schwerpunkt dieser Arbeit stellte die Entwicklung von neuartigen *Tools* zur Verbesserung der Sequenzierung von hmdC in DNA dar. Das in der oxidativen Bisulfit-Sequenzierung (oxBS-seq) genutzte K₂Cr₂O₇, bewirkt meist eine starke Zersetzung der zu sequenzierenden DNA. In dieser Arbeit wurden daher neue und wasserlösliche Derivate des milden Oxidationsmittels IBX entwickelt und synthetisiert, um das K₂Cr₂O₇ in der oxBS-Seq zu ersetzen. Die erfolgreich synthetisierten Derivate *p*-cIBX **22**, IBX-Amid **27** und PEG⁺-IBX **39** wurden anhand von hmdC im Oligonukleotid-Kontext getestet. Es konnte gezeigt werden, dass sich wasserlösliches IBX zur Oxidation von hmdC in Oligonukleotiden (prinzipiell) eignet. Die erreichte Umsatzrate von 75-80 % ist vielversprechend. Bei einer anschließenden Anreicherung der fdC haltigen DNA könnten damit erste Sequenzierungsversuche unternommen werden.

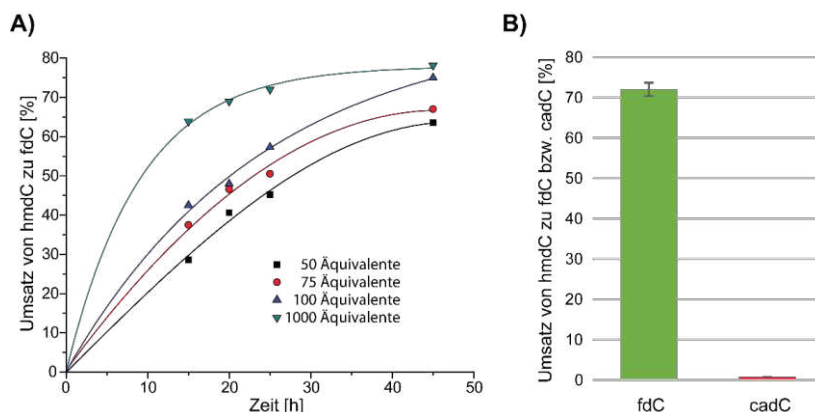


Abbildung II: A) Umsatz von hmdC zum fdC Oligonukleotid, mit 50 Äq *p*-cIBX **22** (schwarzes Quadrat), 75 Äq (roter Kreis), 100 Äq (blaues Dreieck) und 1000 Äq (türkises Dreieck) bzgl. hmdC. Aufgrund einer besseren Übersichtlichkeit wurden Trendlinien eingefügt. B) Menge an fdC und cadC im Vergleich zu hmdC ermittelt mittels quantitativer LC-MS/MS Messung. (siehe Abbildung 52)

Des Weiteren sollte die simultane BS-Seq von mdC, hmdC, fdC und cadC auf doppelsträngiger DNA verbessert werden. Die Anwendung von BS-Seq auf doppelsträngiger DNA wird durch den Einsatz von *Hairpin* Oligonukleotiden möglich. Kürzlich wurde diese Methode auf mdC und hmdC ausgeweitet. Jedoch fehlt bei der verwendeten Methode die Möglichkeit, auch fdC und cadC einzubinden. Daher wurden in dieser Arbeit teilrandomisierte *Hairpin* Oligonukleotide mit allen vier epigenetischen Nukleobasen (mdC, hmdC, fdC und cadC) erfolgreich synthetisiert. Mit diesen Oligonukleotiden ist es nun möglich die jeweilige spezifische Umwandlungsmethode auf ihre Effektivität hin zu überprüfen.

Der zweite Teil dieser Arbeit behandelt die mögliche enzymatische Entfernung von RNA-Modifikationen. Es wurden dabei Erkenntnisse zur Demethylierung der mRNA-Modifikation m⁶A mit Hilfe der Demethylase ALKBH5 erlangt. Dafür wurden unterschiedliche Sequenzmotive mit

m^6A und mit der anspruchsvolleren rRNA-Modifikation m^6_2A mit ALKBH5 umgesetzt. Die durchgeführten Experimente konnten zeigen, dass ALKBH5 die Modifikation m^6A in einem weitaus größeres Sequenzspektrum demethylieren kann, als bisher angenommen. Zudem konnten gezeigt werden, dass ALKBH5 die Modifikation m^6_2A , sowohl im rRNA-Kontext, als auch in kleinen Oligonukleotiden, demethylieren kann. Im konservierten Sequenzkontext von m^6_2A konnten unterschiedliche Intermediate von der einfach demethylierten bis zur vollständig demethylierten Spezies beobachtet werden.

Neben der abundanten Modifikation m^6A befasst sich die vorliegende Dissertation mit der bisher noch wenig erforschten RNA Modifikation ms^2A . Es konnten wertvolle Hinweise über dessen Biosynthese gesammelt werden. Hierfür wurde ein interner Standard von ms^2A synthetisiert, um mit Hilfe der im Arbeitskreis *Carell* etablierten LC-MS/MS-Methodik die Verteilung von ms^2A in unterschiedlichen Spezies zu untersuchen. Dabei konnte ms^2A , übereinstimmend zur Literatur, in menschlichen HAP1-Zellen nachgewiesen werden. Es konnte gezeigt werden, dass diese zum Teil auf die Dealkylierung der tRNA-Hypermodifikation ms^2i^6A zurückzuführen sind.

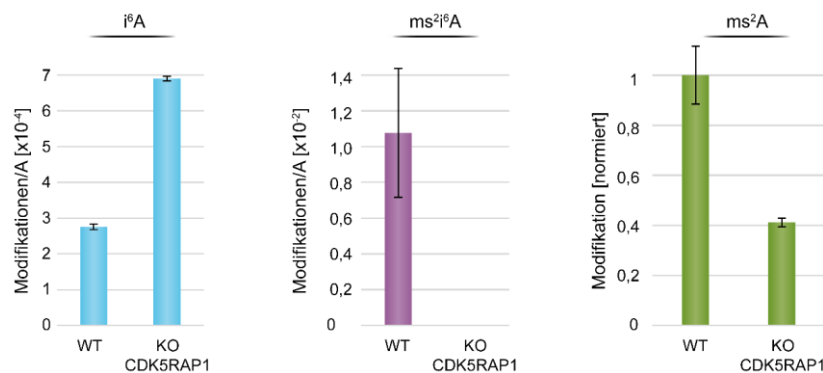


Abbildung III: Quantifizierungsdaten der Modifikationen i^6A , ms^2i^6A und ms^2A , von isolierter gesamt-RNA aus einem HAP1 Wildtyp und dessen CDK5RAP1 Knockout (Enzym zur Biosynthese von ms^2i^6A). Die i^6A Daten (blau) und ms^2i^6A Daten (lila) wurden im Verhältnis zu der Menge an Adenosin aufgetragen. Die Modifikation ms^2A (grün) wurde normiert aufgetragen. (siehe Abbildung 66)

Summary

Multicellular organisms share the same sequence of information, stored as DNA in the cell nucleus. From a single common progenitor, the embryonic stem cells, highly specialized cell types develop. These cell types differ regarding their composition and have different functions. This means that in addition to the base sequence, there must be a further level of information, which must be interpreted individually by the different cell types and which represents a control of gene expression. The DNA modification mdC is an example of this regulatory information level. With the cytosine modifications hmdC, fdC and cadC, three further nucleobases of DNA were discovered in 2009 and 2011. In addition to quantification and in order to better understand their important biological role, the question of their exact positioning in genomic DNA arises. To this end, ways of sequencing these epigenetic nucleobases are still being developed and existing ones are improved. Sequencing methods based on antibody enrichment (which is also very interesting for RNA modifications) or involving chemical derivatisation for example sodium bisulphite (bisulphite sequencing), however, offer great potential for optimisation.

In this dissertation, a novel method for the production of monoclonal antibodies under mild reaction conditions was established. With this method it should now be possible to generate antibodies, not only against sensitive modifications of the base, but also antibodies against modifications of ribose. Haptens were developed for this purpose, which were coupled to carrier proteins (OVA and cBSA) by means of EDC coupling. In addition, ELISA screening was optimized through the use of specially synthesized modified oligonucleotides. With the method developed here, monoclonal antibodies against fdC could be produced.

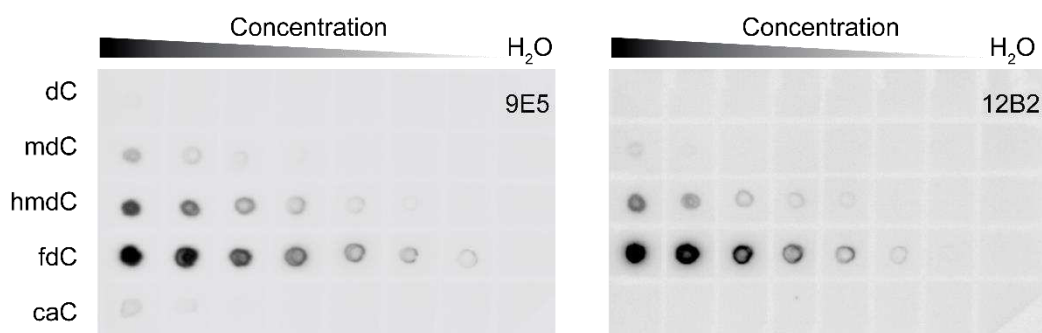


Figure I: Dot blots of the monoclonal antibodies (9E5 and 12B2) against the modification fdC, developed in this thesis. The antibodies were generated using the novel method and in collaboration with the *Helmholtz Zentrum München*. The cross-reactivity with respect to dC and cadC, of the fdC antibodies generated here is significantly lower compared to the commercial product.

These possessed specificities comparable to those of the commercially available polyclonal fdC antibody from Active Motif. In addition, the cross-reactivity between dC and cadC is significantly lower compared to the commercial product.

A further focus of this work was the development of novel tools to improve the sequencing of hmdC in DNA. The KRuO_4 used in oxidative bisulfite sequencing (oxBS seq) usually causes a strong

decomposition of the DNA. In this work, new and water-soluble derivatives of the mild oxidizing agent IBX were developed and synthesized to replace KRuO_4 in oxBS-seq. The successfully synthesized derivatives p-cIBX **22**, IBX-Amid **27** and PEG^+ -IBX **39** were tested with hmdC-containing oligonucleotides. It was shown that water soluble IBX is suitable for oxidation of hmdC in oligonucleotides (in principle). The achieved turnover rate of 75-80 % is very promising. In a subsequent enrichment of the DNA containing fdC, first sequencing attempts could be undertaken.

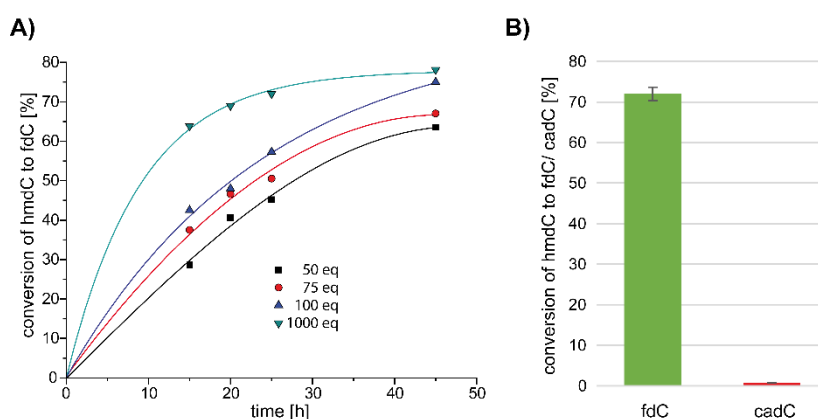


Figure II: A) Turnover from hmdC to fdC in an oligonucleotide context, with 50 eq p-cIBX **22** (black square), 75 eq (red circle), 100 eq (blue triangle) and 1000 eq (turquoise triangle) for hmdC. Trendlines have been added for a better overview. B) Quantity of fdC and cadC compared to hmdC determined by quantitative LC-MS/MS measurement.

Furthermore, the simultaneous BS-Seq of mdC, hmdC, fdC and cadC on double-stranded DNA needs to be improved. The application of BS-Seq on double-stranded DNA is already possible by the use of Hairpin oligonucleotides. Recently this method has been extended to mdC and hmdC. However, the method lacks the ability to integrate fdC and cadC. Therefore, in this thesis, partially randomized hairpin oligonucleotides were successfully synthesized with the four epigenetic nucleobases. With these oligonucleotides it is now possible to investigate the dynamic epigenetic patterns during cell aging, cell division or cell differentiation.

The second part of this work deals with the possible enzymatic removal of RNA modifications. First, the demethylation of the mRNA modification m^6A with the demethylase ALKBH5 was investigated. To this end, different sequence motifs with m^6A and with the more hindered rRNA modification m^6_2A with ALKBH5 were implemented. The experiments carried out showed that ALKBH5 can demethylate the modification m^6A in a much larger sequence spectrum than previously assumed. In addition, ALKBH5 has been shown to demethylate m^6_2A both in the rRNA context and in small oligonucleotides. In the conserved sequence context of m^6_2A , different intermediates could be observed from the simple demethylated to the fully demethylated species.

In addition to the abundant modification m^6A , this dissertation deals with the RNA modification ms^2A , which has so far escaped attention. Valuable insight into the biosynthesis could be collected. In order

Summary

to investigate the distribution of ms^2A in different species using the LC-MS/MS methodology established in the Carell group, an internal standard of ms^2A was synthesized. According to the literature, ms^2A is detected in human cells. It was also shown that ms^2A is partly a dealkylation product of the tRNA hypermodification $\text{ms}^2\text{i}^6\text{A}$.

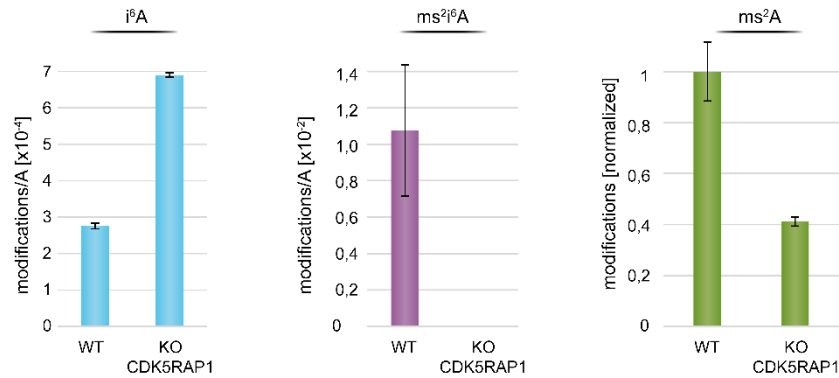


Figure III: Quantification data of the modifications i^6A , $\text{ms}^2\text{i}^6\text{A}$ and ms^2A , of isolated total RNA from a HAP1 WT and its CDK5RAP1 KO (enzyme for biosynthesis of $\text{ms}^2\text{i}^6\text{A}$). The i^6A data (blue) and $\text{ms}^2\text{i}^6\text{A}$ data (purple) were plotted in relation to the amount of adenosine. The modification ms^2A (green) was applied in a normalized manner.

1. Einleitung

1.1. Grundlagen des Aufbaus der DNA

Die Entdeckung der DNA in Zellkernen von Leukozyten durch *Miescher* 1869, stellte den Startpunkt des modernen Verständnisses eines vererbaren Speichers von genetischen Informationen dar.^[1] Jedoch erst ca. 50 Jahre später wurden erste Strukturmodelle des Makromoleküls DNA diskutiert. Anhand von Experimenten mit Hefe-DNA, wurde von *Levene* ein aus Mononukleotid-Bausteinen bestehendes Phosphat-Desoxyribose-Ketten Modell vorgeschlagen.^[2] Die Arbeit von *Avery et al.*^[3] und später *Hershey et al.*^[4] konnten beweisen, dass nicht wie bisher gedacht die Proteine, sondern die DNA zur Weitergabe von Erbinformationen dient. Durch die Erforschung der Verhältnisse der vier kanonischen Basen Adenin (A), Cytosin (C), Guanin (G) und Thymin (T) wurde zudem klar, dass diese A/T und G/C Paare bilden.^[5] Ein weiterer Meilenstein wurde durch die ersten Röntgenaufnahmen eines DNA-Doppelstrangs durch *Wilkins et al.*^[6] und *Franklin et al.*^[7] erreicht. Die gleichzeitige, auf diesen Daten gestützte Aufklärung der Doppelhelixstruktur durch *Watson* und *Crick*,^[8] ermöglichte das bis heute gültige Modell der Basenpaarung. Diese Struktur wird vermittelt durch Wasserstoffbrücken zwischen je einem Purin und einem Pyrimidin (Abbildung 1).^[9]

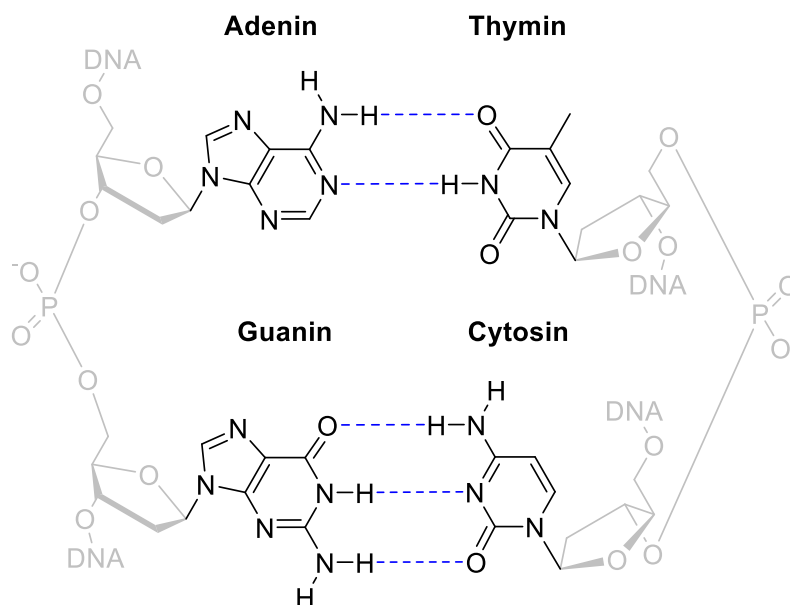


Abbildung 1: Paarung der kanonischen Nukleobasen, nach dem Vorbild von *Watson* und *Crick* in der doppelsträngigen DNA Helix. Die gestrichelten blauen Linien stellen die ausgebildeten Wasserstoffbrücken zwischen den paarenden Basen dar. Dabei wird ersichtlich, dass Adenin und Thymin zwei und Guanin und Cytosin drei Wasserstoffbrücken ausbilden können.

Die Idee eines Replikationsmechanismus der DNA und damit eine mögliche Weitergabe von genetischen Informationen, wurde durch die Entdeckung der DNA Polymerase gestützt.^[10] *Meselson* und *Stahl* konnten kurz darauf experimentell den semikonservativen Charakter der Replikation der DNA beweisen.^[11] Das Modell der Informationsübertragung von der DNA hin zu den Proteinen über

ein Triplett Codon,^[12] das aus diesen grundlegenden Entdeckungen entstand, besteht in den Grundzügen bis heute.^[13]

1.2. Epigenetik

Damit sich aus dem relativ starren genetischen Code, zum einen die beobachtete Vielzahl an hoch spezialisierten Zelltypen entwickeln kann und zum anderen eine ausdifferenzierte Zelle auf Umweltveränderungen reagieren kann, muss es für die Zelle eine Möglichkeit geben in die Genregulation einzugreifen. Da diese Informationen teilweise vererbbar sind, muss es im Genom zwingend eine weitere, eine epigenetische, Ebene geben.^[14]

Schon vor der Strukturaufklärung der DNA wurde der Begriff Epigenetik von *Waddington* ins Leben gerufen, um das Studium des Entwicklungsprozesses vom Genotyp (Erbbild) zum Phänotyp (Erscheinungsbild) zu beschreiben.^[15, 16] Später wurde der Begriff verwendet, um die Untersuchung von „stabilen Veränderungen der Genexpression während der Zellentwicklung und Zellproliferation“ zu beschreiben.^[17] Die wesentlichsten Träger von epigenetischen Informationen werden mit den DNA-Methylierungen, den Komponenten des Heterochromatins (Beispielsweise Histone), den *polycomb* Proteinen und einigen der *noncoding* RNAs beschrieben.^[14] Im Weiteren soll genauer auf die biologische und biochemische Rolle der 5-Cytidin-Modifikationen in DNA (Abbildung 2) eingegangen werden.

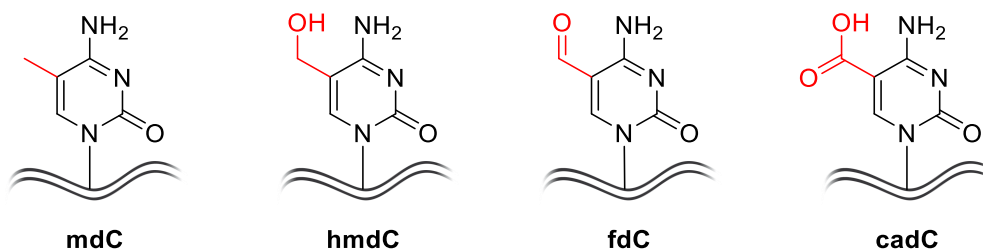


Abbildung 2: Die Darstellung der vier epigenetischen Basen der DNA, welche von Cytidin abgeleitet werden können: 5-Methyl-2'-desoxycytidin (mdC), 5-Hydroxymethyl-2'-desoxycytidin (hmdC), 5-Formyl-2'-desoxycytidin (fdC) und 5-Carboxymethyl-2'-desoxycytidin (cadC).

1.3. Die epigenetischen Basen in der DNA

1.3.1. 5-Methyl-2'-desoxycytidin

Schon früh wurde ein weiterer Bestandteil der DNA, eine Modifizierung des Cytosins, entdeckt.^[18] Das Vorkommen des 5-Methyl-2'-desoxycytidins (mdC), mit einer Methylgruppe an der Basenposition 5, wurde aber erst 1950 von Wyatt in DNA bestätigt.^[19] Das mdC gilt heute, auch aufgrund seiner Abundanz (ca. 4% aller Cytosine sind methyliert), als fünfte DNA Base.^[20, 21, 22] Die Methylmodifikation mdC wird in Eukaryoten durch die drei DNA-Methyltransferasen Dnmt1, Dnmt3a und Dnmt3b, postreplikativ eingeführt. Dnmt1 ist, zusammen mit dem *ubiquitin-like-containing PHD and RING finger domains protein 1* (Uhrf1),^[23] dafür zuständig, während der Replikation der DNA,

eine symmetrische Methylierung im CpG-Kontext aufrecht zu erhalten, indem es den neu synthetisierten Strang an den entsprechenden Positionen methyliert.^[24] Dnmt3a und Dnmt3b, sind hingegen größtenteils für *de novo* Methylierungen, also von bisher nicht methylierten CpG-Dinukleotiden, zuständig.^[25, 26] Aber auch Dnmt1 ist in die *de novo* Methylierungen involviert.^[27] Über eine epigenetische bzw. regulatorische Funktion von mdC wurde schon früh durch unterschiedliche Gruppen diskutiert.^[28, 29, 30] Schon bald konnte die Modifikation, gestützt durch Experimente, mit der Reduktion bzw. Unterdrückung von transkriptionaler Aktivität in Verbindung gebracht werden.^[31, 32, 33] Heute wissen wir, dass mdC zudem an der X-Chromosom-Inaktivierung^[34], sowie der Unterdrückung von Transposons^[35] beteiligt ist und Einfluss auf die genomische Prägung nimmt.^[36, 37] Jedoch ist es bis jetzt noch nicht endgültig geklärt, wie genau die Transkription durch mdC inhibiert wird.^[38]

Interessanterweise wird während der embryonalen Entwicklung eine starke Neusynthese von mdC beobachtet.^[39, 40] Um die damit verbundene Variabilität dieser Modifikation zu erklären, stellte sich die Frage, wie mdC aktiv aus der DNA entfernt werden kann.

1.3.2. 5-Hydroxy-2'-desoxymethylcytidin

Erst 2009 wurde die Modifikation 5-Hydroxy-2'-desoxymethylcytidin (hmdC), welche lange Zeit als oxidativer Schaden galt^[41], als natürlicher Bestandteil der DNA in Säugetieren bestätigt. Gleichzeitig konnten die, für die aktive Oxidation von mdC zu hmdC zuständigen α -Ketoglutarat abhängigen Dioxygenasen *ten-eleven translocation enzymes* (TET 1-3), identifiziert werden.^[42, 43, 44] Bei Untersuchungen der mdC und hmdC Mengen, während der embryonalen Entwicklung, weisen maternale und paternale Stränge unterschiedliches Verhalten in ihren Methylierungs- bzw. Hydroxymethylierungs-Mustern auf.^[39, 45] Dieses Verhalten könnte mit aktiver Demodifizierung erklärt werden. Es konnte gezeigt werden, dass 5-Hydroxymethyluracil (hmU), das prinzipielle Desaminierungsprodukt von hmC, mittels TDG aus der DNA entfernt wird, um im *base excision repair* (BER) mit einem kanonischen Cytosin ersetzen zu werden.^[46] Jedoch konnte dies in murine embryonale Stammzellen (mESC) nicht bestätigt werden.^[47]

1.3.3. 5-Formyl-2'-desoxycytidin und 5-Carboxy-2'-desoxycytidin

2011 wurden die weiteren oxidativen mdC-Derivate 5-Formyl-2'-desoxycytidin (fdC)^[48] und 5-Carboxy-2'-desoxycytidin (cadC)^[49], als Modifikationen der DNA, beschrieben. Auch die Oxidation von hmdC zu fdC und schlussendlich zu cadC wird dabei durch die TET 1-3 Enzyme katalysiert.^[49] Tatsächlich konnte gezeigt werden, dass letztendlich die Basen fC und caC durch TDG direkt ausgeschnitten und mittels BER mit kanonischem Cytosin ersetzt werden.^[50, 51] Damit war ein Weg zur aktiven Entfernung von mdC aus der DNA gefunden, der bis heute Bestand hat (Abbildung 3).

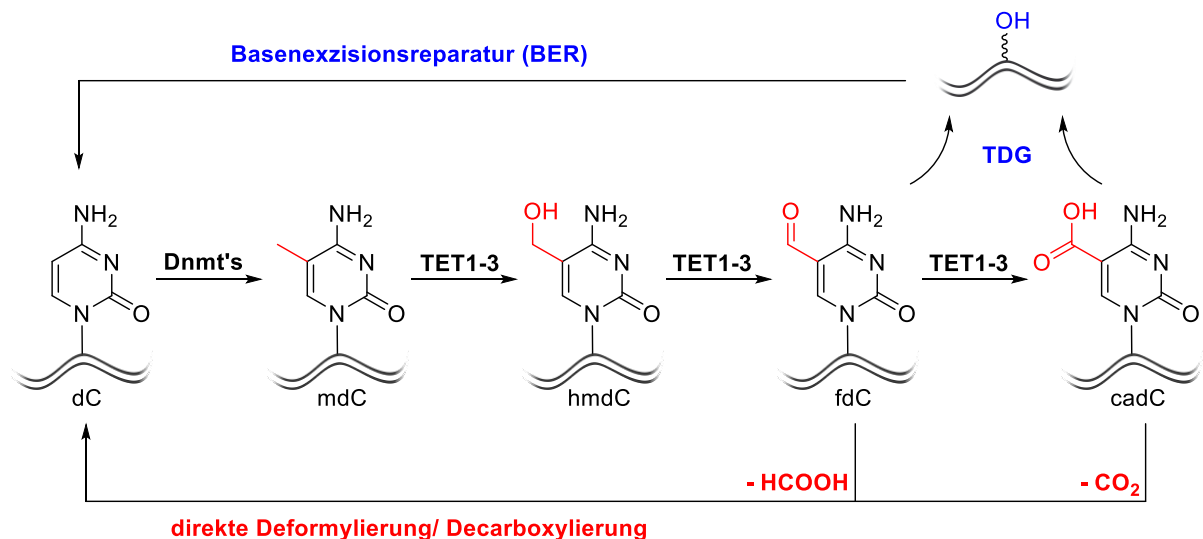


Abbildung 3: Die zwei vermuteten aktiven Demethylierungswege, mit der von den TET Oxygenasen katalysierten Oxidation von mdC zu hmdC, fdC und cadC. Der erste Weg führt über den Ausschnitt der Basen fdC und cadC, durch die Thymin-DNA Glykosylase (TDG). Darauf folgt der *base excision repair* (BER) Prozess (blau), bei dem das ausgeschnittene fdC bzw. cadC durch kanonisches dC ersetzt wird. Der zweite Weg führt über eine direkte Deformylierung von fdC (-HCOOH) bzw. eine direkte Decarboxylierung von cadC (-CO₂), hin zu kanonischem dC (rot). Die Abbildung wurde von Iwan *et al.* adaptiert.^[52]

Interessanterweise konnte aber im Kontrast dazu gezeigt werden, dass in Zygoten die aktive Demethylierung unabhängig von TDG abläuft, aber einen Zusammenhang mit TET3 besitzt.^[53] Das heißt, es muss noch einen weiteren, TDG/BER unabhängigen, Demethylierungsweg geben. Es wird vermutet, dass dies durch einen direkten C-C Bindungsbruch bei fdC bzw. cadC stattfindet. 2017 konnten Iwan *et al.*, mit einer durch Quantifizierungsdaten unterstützten Studie, erstmals eine direkte Deformylierung von fdC bzw. Decarboxylierung von cadC zurück zum Cytosin zeigen.^[52] Jedoch konnte das Enzym, welches den C-C Bindungsbruchs katalysiert, bis jetzt nicht identifiziert werden.

1.3.4. Funktionen der epigenetischen Basen in der DNA

Neben der Funktion der aktiven Demethylierung sind hmdC^[54] und fdC^[55, 56] auch als größtenteils stabile Modifikationen des Genoms etabliert. Die Identifizierung diverser *reader* Proteine der epigenetischen Basen, legt eine mögliche, von der Demethylierung unabhängige, epigenetische Funktion der Basen hmdC, fdC und cadC in der Transkription und Chromatin-Regulation nahe.^[57, 58] Zudem besitzen die epigenetischen Modifikationen einen Einfluss auf die Flexibilität der DNA.^[59] Nicht zuletzt das hohe Vorkommen von hmdC in Gehirngewebe,^[22, 60, 61] bei dem bis zu 1.3% aller Cytosine als hmC Modifikation vorkommen,^[22] bringt diese Modifikation mit der neuronalen Entwicklung in Verbindung.^[62, 63] In einer kürzlich vorgestellten Studie wird hmdC als zentraler Faktor im streng regulierten Prozess der Replikation gesehen und scheint eine inhibitorische Funktion auf die Proliferation auszuüben.^[64] Die Funktionen von fdC und cadC sind hingegen noch größtenteils unbekannt. In einer Studie werden sie mit der Verlangsamung der RNA-Polymerase und der Verringerung ihrer Substratspezifität in Verbindung gebracht.^[65] Zudem konnte gezeigt werden, dass fdC über Ausbildung von Schiff'schen Basen, mit den Lysinen von naheliegenden Histonen in

Verbindung treten und somit regulatorisch in die Transkription eingreifen kann.^[66, 67, 68] Diese Studien zeigen, dass auch die höher oxidierten epigenetischen Basen hmdC, fdC und cadC neben ihrer Rolle in der Demethylierung, eine zentrale Rolle in der Entwicklung der Zelle besitzen.

1.4. Sequenzierung der epigenetischen Basen in der DNA

Mit der Entdeckung von 5-Methyl-2'-desoxycytidin und später dessen oxidierten Derivaten begann, neben der Quantifizierung, die Suche nach seiner genauen Sequenzierung zur Positionsbestimmung in genomischer DNA, um dessen wichtige biologische Rolle genauer zu verstehen.

Die erste Generation der DNA Sequenzierung wurden 1977 von *Maxam* und *Gilbert*,^[69] sowie von *Sanger et. al* ^[70] unabhängig voneinander entwickelt. Die Sanger-Methode nutzt radioaktive Nukleotide, welchen die 3'OH Gruppe fehlt, so dass sie einen Strangabbruch verursachen. Die unterschiedlich langen DNA-Fragmente können über Polyacrylamidgele aufgetrennt werden und lassen einen Schluss auf die Sequenz zu. Diese Methode wurde in den folgenden Jahren optimiert und weitestgehend automatisiert.^[71] Die zweite Generation, oder „*next generation*“ Sequenzierung, wurde durch die Entwicklung des „*sequencing by synthesis*“ Ansatzes möglich.^[72] So konnte nun der Einbau der vier kanonischen Basen direkt abgelesen werden und nicht wie bisher nur die Länge der Fragmente. Dafür werden einzelne DNA-Fragmente an einer Oberfläche gebunden und dort amplifiziert. Die dadurch entstehenden Fragment-Cluster gleicher Sequenz werden mit Hilfe einer DNA-Polymerase in Zyklen von *single base extension* repliziert. Dieser stufenweise Einbau von einer Base pro Zyklus wird durch die verwendeten *fluorescent reversible terminator deoxyribonucleotides* gewährleistet. Neben der reversiblen *terminator* Gruppe, besitzen die Basenbausteine eine reversible fluoreszierende Gruppe, welche einer der vier kanonischen Basen zugeordnet ist. In jedem Zyklus kann somit über die Farbe die jeweilig eingebaute kanonische Base in Echtzeit identifiziert werden (siehe Abbildung 4).^[73]

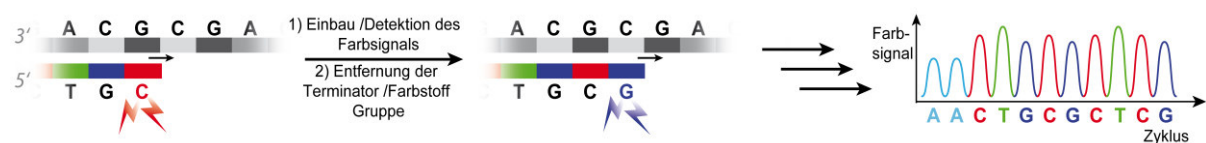


Abbildung 4: Der „*sequencing by synthesis*“ Ansatz zur DNA Sequenzierung. Die fragmentierte genomische DNA (graue Rechtecke) wird amplifiziert und mit Hilfe von *fluorescent reversible terminator deoxyribonucleotides* und einer Polymerase jeweils um eine weitere Base pro Zyklus erweitert (farbige Rechtecke). Ein Detektor empfängt die spezifischen Farbsignale und damit die Sequenz von 5' nach 3'.

Die „*next generation*“ Sequenzierung basiert somit auf der korrekten Basenpaarung durch die DNA-Polymerase. Da die epigenetischen Basen, genau wie Cytosin, über drei Wasserstoffbrücken mit Guanin paaren, können die epigenetischen Basen mit dieser Methode nicht ohne weiteres unterschieden werden und erscheinen als kanonisches Cytosin-Signal. Es besteht daher bis heute eine Notwendigkeit für präzise (*single base resolution*) und sensitive (*single cell sequencing*) Detektionsmethoden der epigenetischen Basen, welche mit der Hochdurchsatzsequenzierung^[74] kompatibel sind und eine hohe Reproduzierbarkeit aufweisen.

Daneben gibt es Bestrebungen, durch eine dritte Generation der Sequenzierungsmethoden, die epigenetischen Basen direkt auszulesen. Hier sind besonders *nanopore sequencing* ^[75] und *single molecule, real-time sequencing* (SMRT-seq) ^[76] zu erwähnen.

1.4.1. Bisulfit Sequenzierung

Prinzipiell basiert der Großteil der Detektionsmethoden für epigenetische Basen auf chemischer Derivatisierung von Cytosin mit Natriumbisulfit (Bisulfit). Die Derivatisierung verändert die Basenpaarung und damit den *read out* in der Sequenzierung. Diese Methode ist nach wie vor der „Goldstandard“ und wird bis heute erfolgreich weiterentwickelt, um immer präzisere Aussagen über die Positionierung der epigenetischen Basen machen zu können.

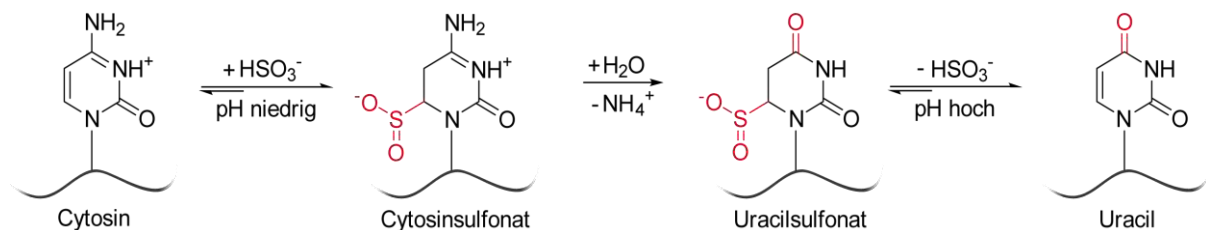


Abbildung 5: Mechanismus der Umsetzung von Cytosin mit Natriumbisulfit und der daraus resultierenden Desaminierung.^[77]

Bei der Behandlung mit Bisulfit im sauren Milieu (pH 5) wird Cytosin zu Cytosinsulfonat umgesetzt. Im Wässrigen hydrolysiert das Cytosinsulfonat, über Desaminierung, zum Uracilsulfonat und kann mittels einer pH Erhöhung zum Uracil reagieren (Abbildung 5).^[77, 78, 79] Unter optimalen Bedingungen kann so das gesamte Cytosin eines Einzelstranges in Uracil umgewandelt werden, was sich in einer Amplifizierung mittels PCR wie Thymin verhält. *Frommer et al.* verband diese chemische Umwandlung mit einer nachfolgenden Sequenzierung, zur „Bisulfit“ Sequenzierung (BS-seq) (Abbildung 6).^[80]

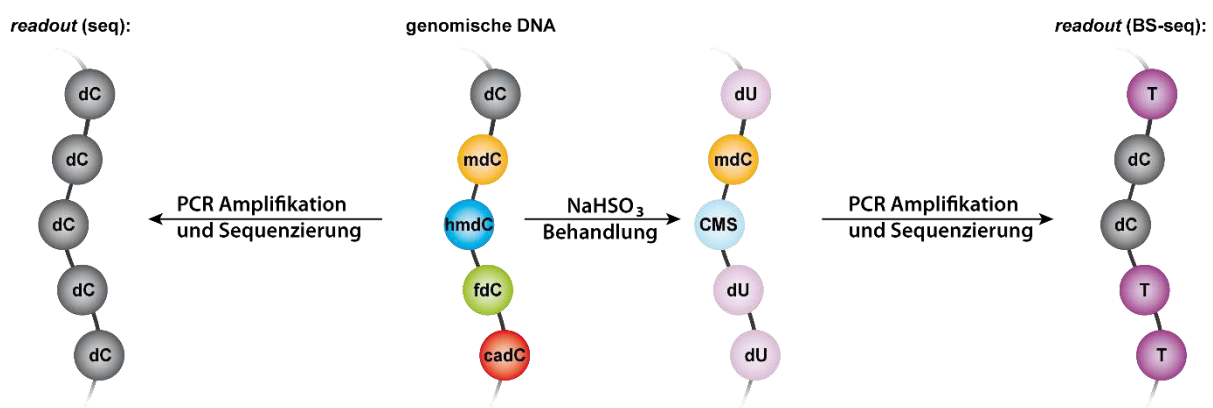


Abbildung 6: Prinzipien der Bisulfit Sequenzierung (BS-seq). Der Vergleich der Sequenzausgabe der normalen Sequenzierung, mit der Sequenzausgabe nach der Behandlung mit Natriumbisulfit zeigt Positionen von mdC und hmdC an. Diese werden nicht desaminiert und werden weiter als dC Signal ausgegeben, wohingegen dC, fdC und hmdC zu dU desaminiert werden und daher ein T Signal ausgeben.

Frommer nutzte dabei die Gegebenheit aus, dass die Methylgruppe des mdC, die Desaminierung entsprechend zum Thymin, verhindert. Durch den Vergleich der Sequenzinformationen mit bzw. ohne Bisulfit-Behandlung, konnten die Positionen von mdC bestimmt werden. Diese Methode wurde noch weiterentwickelt.^[81] So wurden die ersten mdC-Karten erstellt.

Die Entdeckung der Modifikationen hmdC, fdC und cadC in DNA und die Bestätigung, dass diese durch die TET-Enzyme aktiv biosynthetisiert werden, machte es notwendig auch deren Kompatibilität mit der BS-seq Methode zu untersuchen.

1.4.1.1. BS-seq Methoden für hmdC

Wie in Abbildung 7 illustriert, wird hmdC durch die Behandlung mit Bisulfit nicht desaminiert, sondern reagiert sehr effizient zu *cytosine 5-methylenesulfonate* (CMS). Der konkurrierende Angriff an der Position 4 und die damit mögliche Desaminierung ist sehr viel langsamer. CMS verhält sich in der PCR Amplifizierung wie dC und kann daher nicht mehr von mdC unterschieden werden.^[82, 83]

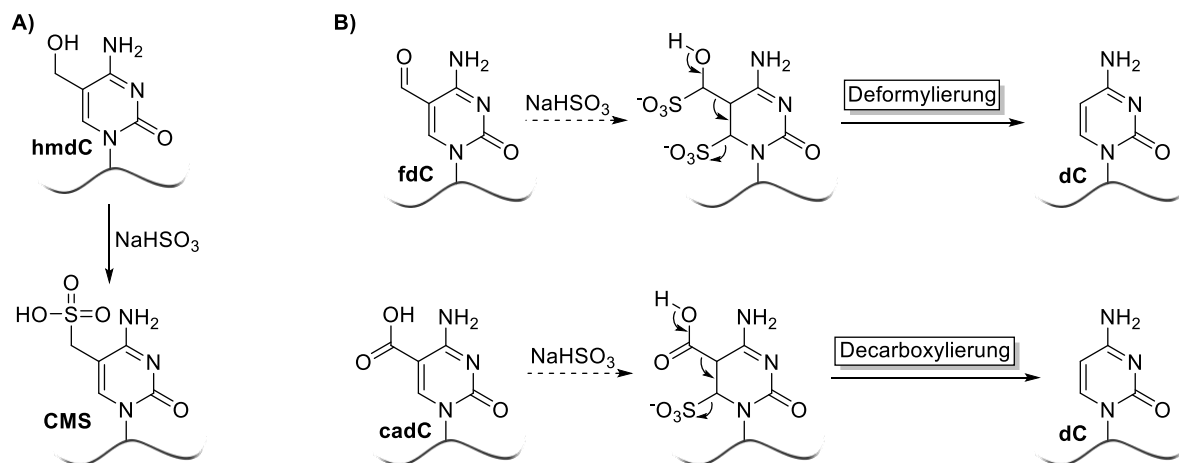


Abbildung 7: Darstellung, wie die epigenetischen Modifikationen hmdC, fdC und cadC mit Bisulfit reagieren A) hmdC wird schnell zum stabilen CMS umgesetzt. B) fdC und cadC werden im Gegensatz dazu über einen noch nicht näher bekannten Mechanismus (gestrichelte Linien) zu Cytosin deformyliert bzw. decarboxyliert. Das so entstehende dC wird im weiteren Verlauf zu dU umgesetzt.^[84]

Bei den bis dahin gesammelten Daten wurden somit neben mdC auch fälschlicherweise die Positionen von hmdC beschrieben (Abbildung 6). Diese neuen Erkenntnisse zeigten die Limitierung des reinen BS-seq auf. Um die Unterscheidung von mdC und hmdC zu ermöglichen, müssen daher zusätzliche Derivatisierungsschritte vor der Behandlung mit Bisulfit durchgeführt werden.

Eine weit verbreitete Methode ist die oxidative Bisulfit Sequenzierung (oxBS-seq).^[85, 86] Hier wird hmdC vor der Behandlung mit Bisulfit spezifisch zu fdC bzw. cadC umgesetzt. Üblicherweise wurde hierfür KRuO₄ genutzt. Die Oxidation von hmdC zu fdC ermöglicht nun eine Reaktion mit Bisulfit und wird im darauffolgenden BS-seq als T gelesen. Der Vergleich der Sequenzen des BS-Seq und des oxBS-Seq zeigen über die T-C Unterschiede, die hmdC Positionen an (Abbildung 8 A). Jedoch sind die Oxidationsbedingungen mittels KRuO₄ relativ harsch und können zu der Zersetzung eines Teils der

DNA führen. Es gibt daher unterschiedliche Ansätze diese Oxidation in milderen Bedingungen durchzuführen, auf die ich später näher eingehe.

Eine weitere Methode zur genomweiten Identifizierung der hmdC Positionen, die etwa zeitgleich zum oxBS-seq entwickelt wurde, ist das *Tet-assisted bisulfite sequencing* (TAB-seq).^[87] Dabei wird hmdC mittels β -Glykosyltransferase zu β -Glykosyl-5-hydroxymethylcytosin (Abbildung 8 B) umgesetzt.

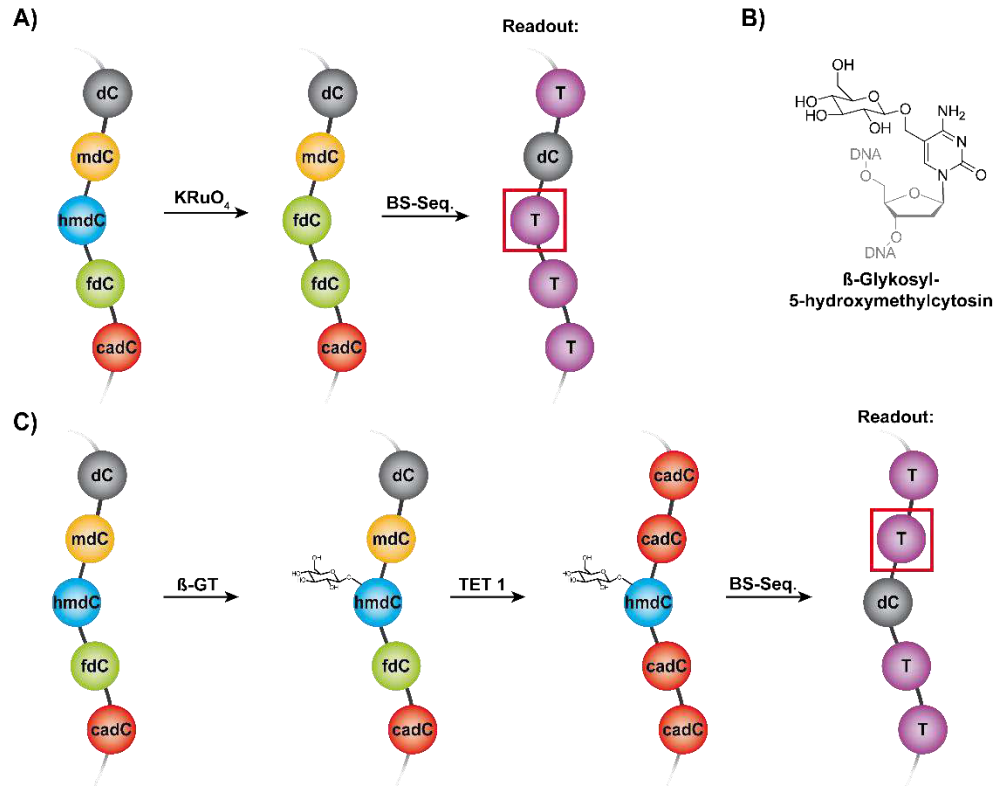


Abbildung 8: A) Illustration der oxidativen Bisulfit Sequenzierung (oxBS-seq). Im ersten Schritt wird hmdC zu fdC oxidiert, um im Folgenden BS-seq, anstatt als dC, als T ausgelesen zu werden. B) Struktur des β -Glykosyl-5-hydroxymethylcytosins. C) Illustration des *Tet-assisted bisulfite sequencing* (TAB-seq). Hier wird hmdC glykosyliert (β -GT = β -Glykosyltransferase) und so vor der Oxidation durch TET 1 geschützt. Im Folgenden BS-seq wird nur noch hmdC als dC ausgelesen. Die roten Vierecke beschreiben die Positionen in der die Unterschiede zum klassischen BS-seq erreicht werden.

Beim TAB-seq wird die DNA im folgenden Schritt zusätzlich mit rekombinantem TET1-Protein behandelt. TET1 oxidiert mdC, zu hmdC und weiter zu fdC und cadC. Die Glykosylierung schützt dabei das hmdC vor der Oxidation durch TET1. Somit werden, bis auf das β -Glykosyl-5-hydroxymethylcytosin, alle epigenetischen Basen zu caC umgesetzt. Bei der folgenden BS-seq wird somit nur noch das geschützte hmdC als dC gelesen und kann von mdC unterschieden werden (Abbildung 8 C).

1.4.1.2. BS-seq Methoden für fdC

Die Modifikationen fdC und cadC können im BS-seq von dC nicht unterschieden werden. Im Gegensatz zu hmdC, werden sie deformyliert bzw. decarboxyliert (Abbildung 7). Das so entstehende dC kann wiederum zu dU desaminiert werden. Somit verhalten sich fdC und cadC während des BS-Seq, wie dC.

Ähnlich zur oxBS-Seq wurde von *Booth et al.* eine Strategie zur Erkennung von fdC im BS-Seq entwickelt, das *reductive bisulfite sequencing* (redBS-seq).^[88] Auch hier wird das unterschiedliche Verhalten von fdC und hmdC im BS-seq genutzt. Dabei wird fdC mit NaBH₄ zu hmdC reduziert. Das hmdC wird durch die Behandlung von Bisulfit zu CMS umgesetzt und wird weiterhin als dC gelesen. Wiederum die Unterschiede des redBS-seq zum Standard BS-seq, zeigen die Positionen von fdC im Genom (Abbildung 9 A).

Eine Strategie zur Sequenzierung von fdC, mit Hilfe chemischer Modifizierung, ist das *5fC chemical modification-assisted bisulfite sequencing* (fCAB-seq).^[89] Beim fCAB-seq wird die Formyl-Gruppe, durch die Bildung einer Schiff'schen Base mit einem Hydroxylamin-Derivat, vor der Bisulfit-vermittelten Desaminierung geschützt. Dadurch wird fdC, wie im redBS-seq Ansatz, in der Sequenzierung als C gelesen (Abbildung 9 B).

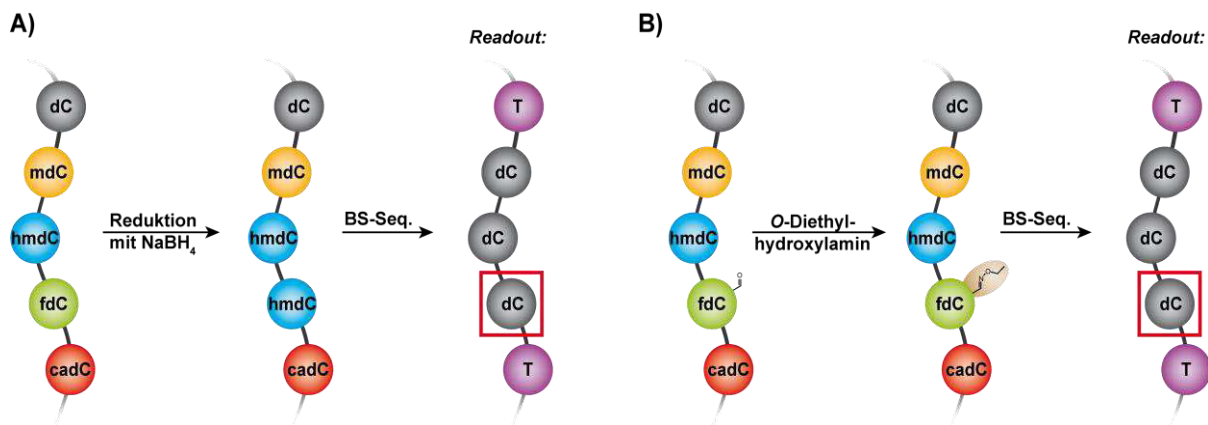


Abbildung 9: A) Illustration des *reductive bisulfite sequencing* (redBS-seq). Im ersten Schritt wird fdC zu hmdC reduziert, um im Folgenden BS-seq, anstatt als T, als dC ausgelesen zu werden. B) Illustration des *5fC chemical modification-assisted bisulfite sequencing* (fCAB-seq). Hier wird fdC mit einem *O*-Diethylhydroxylamin umgesetzt. In dem darauffolgenden BS-seq wird nun auch das geschützte fdC als dC ausgelesen. Die roten Vierecke zeigen die Positionen, welche zum klassischen BS-seq verändert ausgelesen werden.

Statt mit einem einfachen *O*-Diethylhydroxylamin kann fdC auch mit komplizierteren Hydroxylaminen umgesetzt werden und in verschiedenen Bereichen eingesetzt werden. Beispielsweise durch die Verwendung eines Hydroxylamin-Fluorophors. Mittels Fluoreszenzmessungen kann so die fdC Konzentration des DNA Samples bestimmt werden und eine Aussage von fdC-fdC Symmetrien gemacht werden.^[90] Die Verwendung eines Hydroxylamin-Biotins hingegen ermöglicht die Anreicherung von fdC DNA Fragmenten.^[91]

1.4.1.3. BS-seq Methoden für cadC

Das *methylase-assisted bisulfite sequencing* (MAB-seq) ist eine enzymatische Methode zur Bisulfit basierten fdC Sequenzierung.^[92, 93] Mit der *S*-Adenosylmethionin (SAM) abhängigen CpG Methyltransferase (M.SssI) können alle CpG Positionen eines Doppelstrangs methyliert werden. Das so entstehende mdC wird nicht mehr durch Bisulfit umgesetzt. Nach der Bisulfit-Behandlung werden nur noch fdC und cadC als T gelesen. Jedoch kann so nicht zwischen fdC und cadC unterschieden

werden. Diese Methode wurde daher zur *5caC methylase-assisted bisulfite sequencing* (caMAB-seq) weiterentwickelt. Hier wird vor der Bisulfit-Behandlung das fdC mittels NaBH_4 zu hmdC reduziert und somit nur noch cadC als T gelesen (Abbildung 10).^[92, 94]

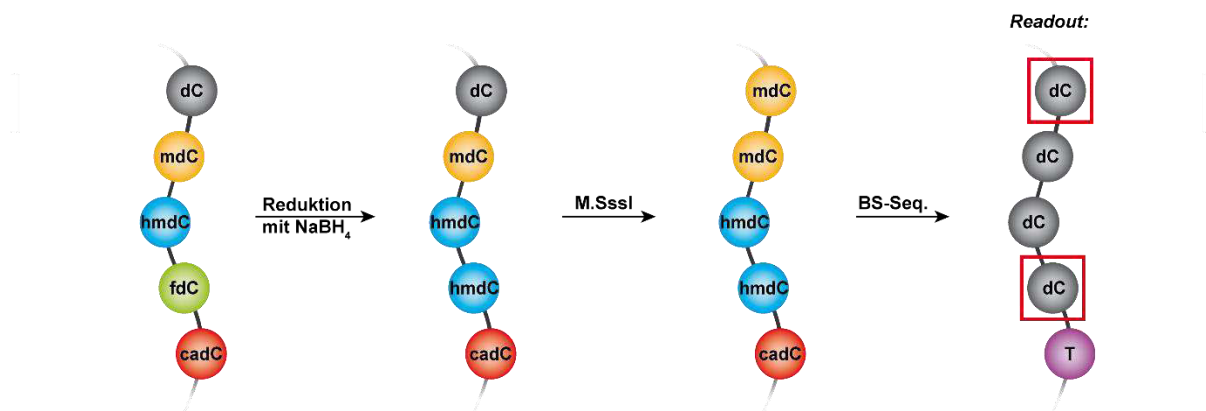


Abbildung 10: Illustration des *methylase-assisted bisulfite sequencing* (MAB-seq). Im ersten Schritt wird fdC zu hmdC reduziert. Durch die Methyltransferase M.SssI wird das gesamte dC (im CpG Kontext) zu mdC umgesetzt. Damit kann in der Bisulfit-Behandlung nur noch cadC decarboxyliert und deasaminiert werden. Was einen direkten cadC readout möglich macht. Die roten Vierecke beschreiben die Positionen, in der Unterschiede zum klassischen BS-seq erreicht werden.

Durch die Kombination der hier beschriebenen Methoden können die epigenetischen Basen in Einzelbasen-Auflösung kartiert werden. Jedoch hat die Bisulfit-Behandlung einen entscheidenden Nachteil, sie zersetzt einen Großteil der zu untersuchenden DNA (bis zu 90%). Die dabei entstehenden drei- oder zwei-Basen-Fragmente können zur späteren Annotierung der Sequenzen nicht mehr verwendet werden^[95, 96] Bei der Untersuchung von verschiedenen Methoden zur *library preparation* wurde festgestellt, dass die gewonnen absoluten und relativen Werte der Methylierungslevel stark variieren.^[96] Daher ist es wichtig, dass die Methoden weiterhin verbessert werden.

1.4.1.4. Hairpin Sequenzierung

Die Vererbung des Methylierungsmusters ist eine zentrale Eigenschaft des epigenetischen Prozesses in DNA, da es wichtige regulatorische Aufgaben übernimmt.^[17, 57, 97] Es ist deshalb von großer Bedeutung auch die Übertragung des Methylierungsmusters auf den Tochterstrang nach der Replikation auf eventuell auftretende Asymmetrien zu untersuchen. Bei der Entwicklung des BS-Seq wurde beobachtet, dass dieses nur auf einzelsträngiger DNA (ssDNA) effektiv funktioniert. Bei der Behandlung von doppelsträngiger DNA (dsDNA) werden im Kontrast nur 0,1% der Cytosine umgesetzt.^[81] Die Sequenzinformation und das Methylierungsmuster von dsDNA muss damit in zwei verschiedenen Reaktionsansätzen mit verschiedenen *Primern* durchgeführt werden.^[98] Durch die erforderliche Denaturierung können zusammengehörige *sense* und *antisense* Stränge nicht mehr eindeutig als diese identifiziert werden und die Informationen über eine Hemimethylierung geht damit verloren. Dies wird durch die starke DNA-Degradierung, während der harschen Bedingungen der Bisulfit-Behandlung und der Abhängigkeit zur vollständigen Umsetzung von Cytosin zu Thymin, weiter verstärkt.^[95]

Durch die Entwicklung des *hairpin bisulfite sequencing* (HP-BS) von *Laird et al.* konnte dieses Problem geschickt umgangen werden.^[99] In dieser Methode werden die *sense* und *antisense* Stränge kovalent über einen Hairpin miteinander verbunden. Dafür wird die dsDNA mit einer ausgewählten Restriktions-Endonukleasen geschnitten. Der dadurch entstehende spezifische Überhang kann durch ein gezielt entworfenes Hairpin Oligonukleotid gebunden werden. Durch die Ligation der *Hairpins* mit der dsDNA sind der *sense* und der *antisense* Strang während der Bisulfit-Sequenzierung kovalent verbunden und können als Einheit abgelesen werden.^[99]

1.4.2. Bisulfit freie Sequenzierung

1.4.2.1. Sequenzierung mit Hilfe von Endonukleasen

Mit der *HpaII tiny fragment enrichment by ligation-mediated PCR* (HEPL)^[100] Analyse und deren Kombination mit *next-generation sequencing* (Methyl-seq)^[101] wurden zwei Methoden entwickelt, welche mittels Endonukleasen erstmals genomweite Methylierungskarten möglich machten. Diese Methoden beruhen auf dem Verdau mit *HpaII* und *MspI*. Der Unterschied dieser Endonukleasen liegt darin, dass *HpaII* die Sequenz CCGG nicht schneiden kann wenn sie methyliert bzw. hydroxymethyliert vorliegt, *MspI* hingegen schon.^[100] Durch den Vergleich der Fragmente, kann auf die Position von mdC bzw. hmdC geschlossen werden (siehe Abbildung 11).

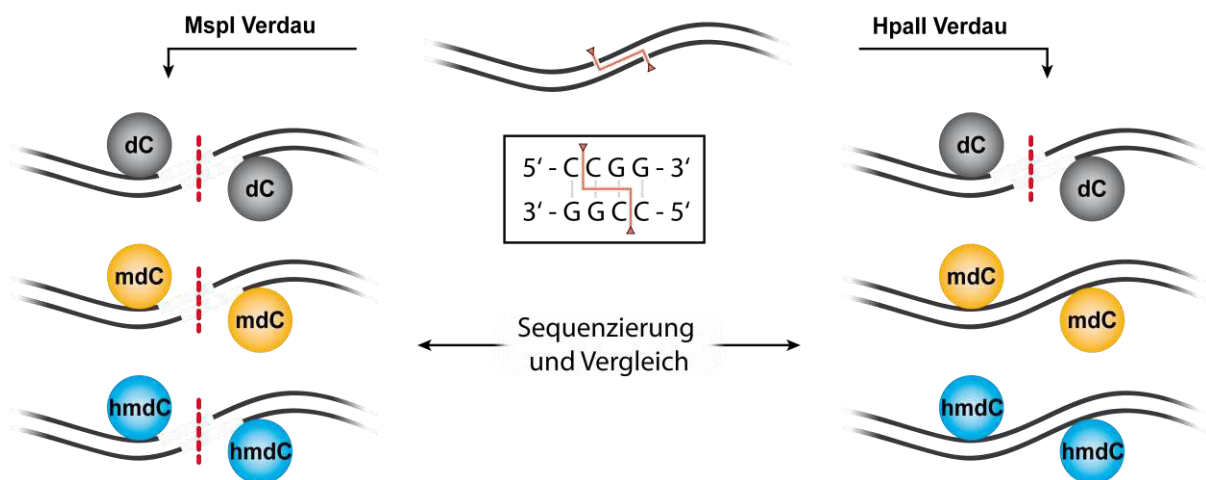


Abbildung 11: Sequenzierung von mdC und hmdC mit Hilfe von Endonukleasen. Die Endonukleasen *MspI* und *HpaII* verdauen dieselbe Restriktionsschnittstelle CCGG (schwarzer Kasten). *HpaII* kann, im Kontrast zu *MspI*, die Restriktionsschnittstelle nicht schneiden, wenn sie mdC oder hmdC beinhaltet. Durch den Vergleich der jeweils erzeugten Fragmente kann auf mdC bzw. hmdC geschlossen werden.

Schon kurz nachdem hmdC als DNA Modifikation bestätigt wurde mit *PvuRts11*, auch hierfür eine spezifische Endonuklease gefunden.^[102]

1.4.2.2. Chemische Derivatisierung und Anreicherung

Ein weiterer großer Bereich der Bisulfit-freien Sequenzierungsmethoden basiert auf weiteren chemischen Derivatisierungen, verbunden mit Anreicherung und/oder dem blockieren der DNA Polymerase in *primer extension assays*.^[103] *Song et al.* entwickelte für die selektive chemische

Derivatisierung von hmC (hmC-seal), eine sehr vielseitige Methode, welche hinsichtlich ihrer Sensitivität immer noch verbessert wird.^[104, 105]

Eine weitere direkte und selektive Derivatisierung von fdC, mit Malonsäuredinitril, findet im *chemical-labeling-enabled C-to-T conversion sequencing* (CLEVER-seq) Anwendung.^[106] Durch die Umsetzung mit Malonsäuredinitril wird das modifizierte fdC in der folgenden Amplifizierung als T ausgelesen. Mit dieser Methode kann eine basengenaue Auflösung von fdC im gesamten Genom einzelner Zellen erreicht werden. Jedoch hängt auch hier die Qualität der Daten direkt mit der Vollständigkeit der fdC-Malonsäuredinitril Umsetzung zusammen.

Eine erst 2019 veröffentlichte Methode kann fdC und cadC, mit Hilfe von Pyridinboran, in Dihydrouracil (DHU) umwandeln.^[107] Das Uracil-Derivat DHU wird bei der darauffolgenden Sequenzierung folglich als Uracil bzw. Thymin gelesen (siehe Abbildung 12).

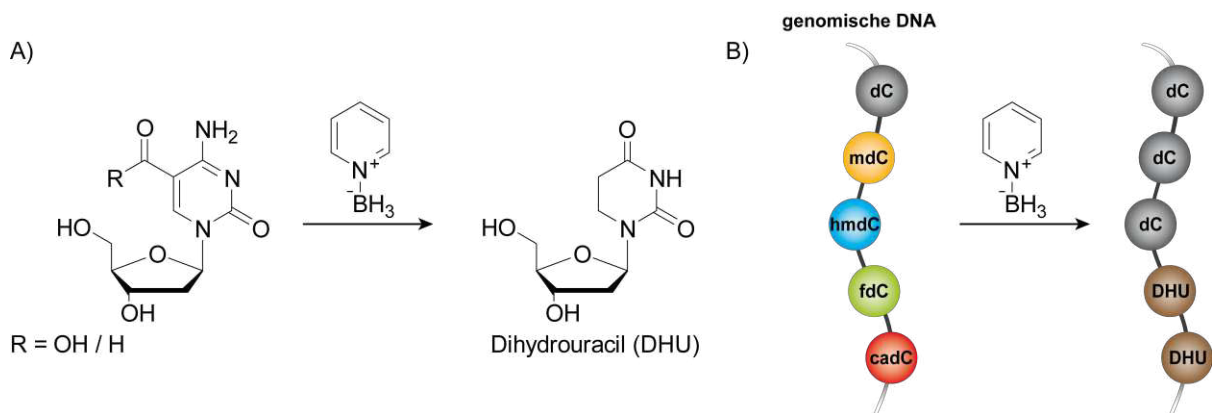


Abbildung 12: A) Umsetzung von cadC bzw. fdC mit Pyridin-Boran zu Dihydrouracil (DHU). B) Prinzipien des *pyridine borane sequencing*. cadC und fdC werden von Pyridin-Boran zu DHU umgesetzt und als T Signal in der darauf folgenden Sequenzierung ausgegeben, wohingegen dC, mdC und hmdC nicht verändert werden und damit als C Signal erscheinen.^[107]

Im Stil der BS-seq, können mit Hilfe von selektiver chemischer Umwandlung der epigenetischen Basen (oxBS oder redBS) und dem Vergleich der Ergebnisse auf die basengenaue Position dieser geschlossen werden (z.B. *TET-assisted pyridine borane sequencing* (TAPS)). Durch die vergleichsweise milden Reaktionsbedingungen soll die direkte Sequenzierung von bis zu 10 Kilobasen lange DNA Fragmente möglich sein.^[107] Ob sich damit, im Vergleich zu den auf Bisulfit basierenden Methoden, eine bessere Reproduzierbarkeit der Sequenzierungsergebnisse erreichen lässt, bleibt abzuwarten.

1.4.2.3. Sequenzierung mit Hilfe von Antikörpern

Der letzte große Bereich innerhalb der Sequenzierbemühungen, ist die Immunpräzipitation mittels spezifischer Antikörper. Ein Beispiel ist das *methylated DNA immunoprecipitation sequencing* (MeDIP-Seq).^[108] Bei dieser Methode wird die genomische DNA fragmentiert, meist denaturiert und dann über einen mdC spezifischen Antikörper immunpräzipitiert. Bei der Sequenzierung häufen sich die *reads* bei den mdC angereicherten Positionen und es kann auf die Position von mdC geschlossen werden. Die hmdC Antikörper waren, zu dieser Zeit, nicht spezifisch genug, um mit ihnen ein

vergleichbare Methode zu entwickeln. Das änderte sich durch die Entwicklung eines CMS Antikörpers.^[109] Damit konnte hmdC, nach Behandlung mit Bisulfit und darauf folgender Immunpräzipitation, genomweit sequenziert werden.^[110]

Jedoch wird die Selektivität der bisher verwendeten Antikörper kontrovers diskutiert.^[111] Es gibt Hinweise darauf, dass die verwendeten IgG Antikörper auch kurze, intrinsische und unmodifizierte DNA-Wiederholungen binden könnten.^[111]

1.5. Modifizierte Nukleoside in der RNA

RNA-Modifikationen zeigen eine sehr viel höhere chemische Diversität, als DNA-Modifikationen. Bis heute wurden mehr als 170 Modifikationen der RNA entdeckt und diese sind in allen Domänen zu finden.^[112] Diese können von meist einfachen Methylierungen (z.B. *N*⁶-Methyladenosin (*m*⁶A) oder *N*⁷-Methylguanosin (*m*⁷G))^[113], bis hin zu anspruchsvollen, über viele Schritte aufgebauten, Hypermodifikationen (z.B. Queuosin (Q), Wybutosin (yW) oder *N*⁶-Isopentenyl-2-thiomethyladenosin (*ms*²*i*⁶A)) reichen, die nur noch wenig mit der kanonischen Base, von der sie sich ableiten, zu tun haben (Abbildung 13).^[112, 114]

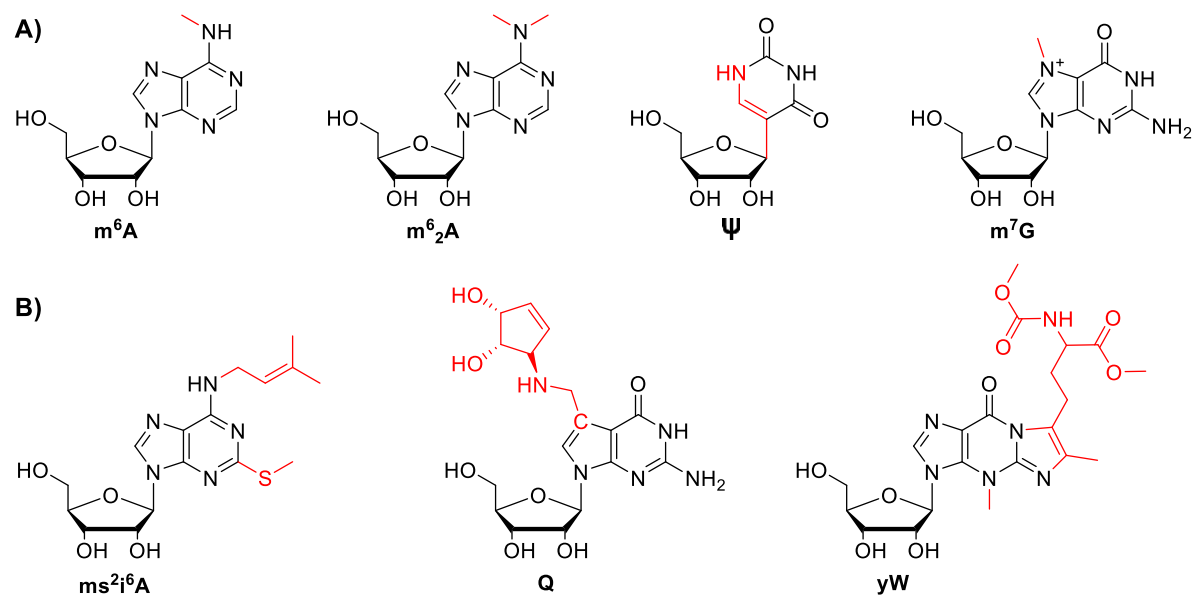


Abbildung 13: A) Schematische Darstellung einer Auswahl an RNA Modifikationen B) Eine Auswahl an Hypermodifikationen, welche hauptsächlich in *transfer* RNA (tRNA) gefunden wurden.

Mit zunehmender Sensitivität der Analysemethoden werden weiterhin neue Modifikationen der RNA entdeckt.^[115, 116, 117] Einige dieser Modifikationen, sind evolutionär konserviert und es wird angenommen, dass diese schon im *last universal common ancestor* (LUCA) existierten.^[118] Dabei findet man Modifikationen in nahezu allen RNA Spezies, in denen sie mit unterschiedliche Funktionsbereichen verknüpft sind.^[119]

Die am stärksten modifizierten RNA Spezies ist die transfer RNA (tRNA). Hier sind auch die meisten hypermodifizierten Basen zu finden.^[120] Die ribosomale RNA (rRNA) ist eine weitere hochmodifizierte

RNA Spezies und besitzt mit beispielsweise N^6,N^6 -Dimethyladenosin (m^6A)^[121] stark konservierte Modifikationen.

1.5.1. Nukleosid-Modifikationen in mRNA

Lange Zeit führten die im Verhältnis kurzlebige, *messenger* RNA (mRNA), neben den stark modifizierten und abundanten Spezies tRNA und rRNA, ein Schattendasein. In mRNA wurde schon früh die Modifikation m^6A , entdeckt und beschrieben.^[122, 123] Jedoch war diese, trotz ihres Potentials auf die Genexpression Einfluss nehmen zu können, eher unbeachtet. Denn die zur Verfügung stehenden Detektionsmethoden, waren auf die Quantifizierung von verdauter RNA, limitiert.^[124]

Das änderte sich, als m^6A erstmals über *methylated RNA immunoprecipitation sequencing* (MeRIP-seq) transkriptomweit sequenziert werden konnte und dabei in hoher Abundanz in mRNA bestätigt wurde.^[125, 126]

1.5.2. Das Epitranskriptom

Angelehnt an den Begriff Epigenom aus dem Feld der DNA-Epigenetik, entwickelte sich der Begriff Epitranskriptom, im Bereich der dynamischen RNA-Modifikationen.^[126, 127] Das Feld der Epitranskriptomik beschreibt die Erforschung der RNA-Modifikationen und deren funktionellen Einfluss auf die Regulation der Genexpression mittels RNA-Prozessierung. Vergleichbar zur Epigenetik wird durch die post-transkriptionalen Modifikationen die Sequenz der RNA *per se* nicht verändert, sondern beschreibt, wie das Epigenom, eine zusätzliche Ebene der Information.^[124] Kürzlich konnte gezeigt werden, dass die 2'-*O*-Methylierungen der Ribose in mRNA die Genexpression regulieren können.^[128] Der Begriff ist jedoch nach wie vor stark mit der mRNA-Modifikation m^6A verknüpft. Durch den dynamischen Charakter von m^6A , wird davon ausgegangen, dass ein Organismus damit auf externe Umwelteinflüsse reagieren kann.^[129]

1.5.2.1. *Writer* und *reader* von m^6A

Neben der Sequenzierung von m^6A konnten mit den Methyltransferasen *methyltransferase-like 3* und *14* (METTL3 und METTL14), auch die Proteine identifiziert werden, welche den posttranskriptionalen Einbau von m^6A katalysieren.^[130] Ein Heterodimer aus METTL3 und METTL14 bildet einen Komplex mit dem Adaptor-Proteinen *wilms tumor 1-associated protein* (WTAP)^[131] und weiteren Faktoren (KIAA1429,^[132] RBM15/B,^[133] ZFP217^[134, 135] und weitere^[136]). Dieser Komplex kann, mit dem Coenzym *S*-Adenosylmethionin (SAM) als Methyl donor, eine Methylgruppe auf Adenosin übertragen. Während METTL3 die katalytische Aktivität besitzt, ist METTL14 dabei für die Substraterkennung zuständig.^[137]

Schon früh wurden mit den YTHDF1-3 erste Proteine identifiziert, die mit m^6A direkt interagieren.^[125] Neben YTHDF1-3 wurden auch YTHDC1-2^[138, 139] als m^6A *reader* in menschlichen Zellen identifiziert. Diese Proteine besitzen alle eine *YTH521-B homology* (YTH) Domäne, mit der sie m^6A über hydrophobe Wechselwirkungen binden können.^[138, 140] Es konnten aber auch *reader* gefunden werden, die diese Domäne nicht besitzen. Beispiele dafür sind ELAVL1,^[125] HNRNPA2B1,^[141] IGF2BP1,^[142]

HNRNPC^[143] und METTL3,^[144] um nur einige zu benennen. Hierbei findet die Wechselwirkung meist indirekt über strukturelle Veränderungen der RNA statt, hervorgerufen durch m⁶A.^[143, 145] Über diese spezifisch interagierenden Proteine kann m⁶A vielfältige biologische Funktionen beeinflussen. Beispielsweise gibt es Hinweise, dass m⁶A im Kontext von pre-mRNA, über YTHDC1 alternatives Spleißen hervorrufen kann.^[146] Zudem hat m⁶A einen Einfluss auf die Stabilität der mRNA,^[142, 147] beeinflusst die Neuronale Differenzierung über den Zellkern Export von mRNA^[148] und kann regulatorisch in die Genexpression eingreifen,^[149, 150] um nur einige zu nennen.

1.5.2.2. Eraser von m⁶A

Die Entdeckung der katalytischen und damit aktiven Demethylierung von m⁶A in RNA, durch das *fat mass and obesity-associated* (FTO) Protein, vollendete das Bild von m⁶A, als eine dynamische Modifikation des Epitranskriptoms.^[151] Es wird angenommen, dass die α -Ketoglutarat abhängige Dioxygenase FTO die Methylgruppe von m⁶A zu N⁶-Hydroxymethyladenosin (hm⁶A) und weiter zu N⁶-Formyladenosin (f⁶A) oxidiert. Die oxidierten Derivate (hm⁶A und f⁶A) sind unter physiologischen Bedingungen nicht stabil und zerfallen zu Adenosin (Abbildung 14).^[152] Jedoch wird die Demethylierung durch FTO mittlerweile kontrovers diskutiert^[153] und heute eher mit der Demethylierung von m⁶Am zu 2'-O-Methyladenosin (Am) in der mRNA *cap*-Struktur in Verbindung gebracht.^[154, 155] Im Kontrast dazu stehen aktuelle Strukturanalysen des katalytischen Mechanismus von FTO, welche m⁶A als das Hauptsubstrat sehen.^[156] Kurz nach der Entdeckung der Demethylaseaktivität von FTO, wurde mit *α -ketoglutarate-dependent dioxygenase alkB homolog 5* (ALKBH5) eine alternative m⁶A Demethylase entdeckt.^[157] Im Gegensatz zu FTO besitzt ALKBH5 keine Aktivität gegenüber m⁶Am.^[154]

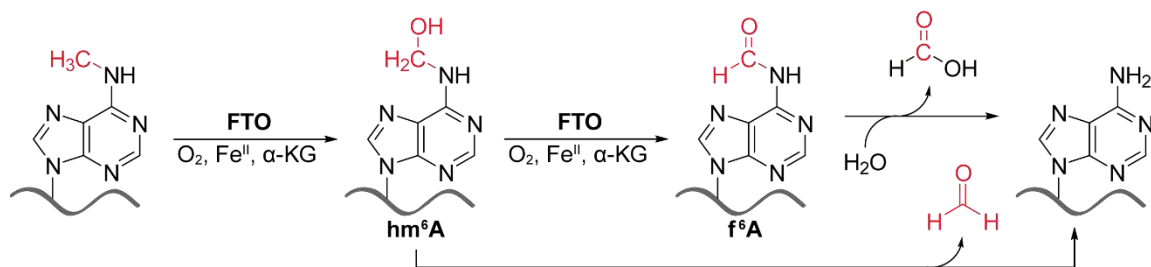


Abbildung 14: Vermuteter Mechanismus zur Demethylierung von m⁶A mittels FTO. Die Methylgruppe des m⁶A wird zum hm⁶A und weiter zum f⁶A oxidiert. Unter physiologischen Bedingungen zerfällt hm⁶A in Ameisensäure und A. Das f⁶A zerfällt mit Wasser in Ameisensäure und A. Die Abbildung wurde von Fu et al. adaptiert.^[152]

ALKBH5 zeigt eine starke Substratspezifität gegenüber m⁶A in einzelsträngiger RNA.^[158, 159] Interessanterweise konnte gezeigt werden, dass ALKBH5 auch an einzelsträngiger DNA arbeiten kann, Doppelstränge sind jedoch aufgrund einer konservierten Disulfidbrücke ausgeschlossen.^[159] Eine Kombination aus *accelerated molecular dynamics* (aMD) Simulation und NMR Strukturanalysen vom humanen ALKBH5, zeigen zudem einen sehr viel dynamischeres Enzym, als bisher, ausgehend von Kristallstrukturen, angenommen. Dabei konnte gezeigt werden, dass ALKBH5 ungeordnete und geordnete Stadien durchläuft, um beispielsweise die Bindungstasche für α -Ketoglutarat zugänglicher

zu machen.^[129] Aber auch bei ALKBH5 bleibt der genaue Demethylierungsmechanismus bis heute unbekannt. Anders als bei FTO scheint ALKBH5 m⁶A aber ohne erkennbare Zwischenstufen zu demethylieren.^[160]

1.5.3. Dealkylierung in tRNA

Die tRNA ist die am stärksten modifizierte RNA-Spezies, jedoch sind die modifizierten Basen in ihrer Abundanz innerhalb der tRNA sehr ungleich verteilt. Es gibt Positionen, an denen keine oder nur wenige Modifikationen nachgewiesen wurden und dann gibt es Positionen die fast ausschließlich mit modifizierten Basen besetzt sind.^[112] Zu diesen nahezu quantitativ modifizierten Bereichen der tRNA gehört die direkt am Anticodon-Triplet angrenzenden Positionen 37. Die Position 37 ist dabei ausschließlich mit Purinen besetzt.^[161] Zudem sind hier die beiden bisher bekannten humanen 2-thiomethylierten Adenosin-Modifikationen lokalisiert: ms²i⁶A und N⁶-Treonylcarbamoyl-2-thiomethyl-adenosin (ms²t⁶A).^[112, 162]

Die Biosynthese der ms²A-Hypermodifikationen wird jeweils durch Enzyme der Radikal-SAM-Familie katalysiert. Für die Biosynthese von ms²i⁶A in Säugetieren ist das Enzym *CDK5 regulatory subunit-associated protein 1* (CDK5RAP1) Verantwortlich.^[163] Es überträgt die Thiomethyl-Gruppe mit Hilfe von zwei [4Fe-4S] Clustern auf die biologische Vorstufe i⁶A.^[164] Die Modifikation ms²t⁶A wird in Säugetieren von *CDK5 regulatory subunit associated protein 1-like 1* (CDKAL1) eingeführt. Auch hier ist das N⁶-modifizierte Adenosin (t⁶A) die biologische Vorstufe.^[162]

Dem weiteren Metabolismus dieser Hypermodifikationen wurde jedoch bisher nur wenig Beachtung geschenkt. Eine in diesem Zusammenhang sehr interessante Thiomethyl-Modifikation ist 2-Thiomethyladenosin (ms²A), welche erstmals 1978 beschrieben wurde. ms²A wurde damals als ein mögliches Abbauprodukt der Modifikationen ms²i⁶A und ms²t⁶A während der RNA Isolation beschrieben.^[165] Jedoch konnte Kellner *et al.* ms²A kürzlich in kleinen RNA-Spezies (70-100 nt) in *E. coli* und in geringerer Konzentration auch im Schwein und im Menschen detektieren.^[115, 166] Dies weist daraufhin, dass es sich um eine Modifikation der tRNA handeln könnte. In *E. coli* gibt es zudem Hinweise darauf, dass ms²A mit dem Abbau von ms²i⁶A zusammenhängen könnte.^[166] Über ms²A ist darüber hinaus nur wenig bekannt und es wird noch nicht in der Datenbank für RNA-Modifikationen *Modomics* geführt.^[112]

2. Aufgabenstellung

Um Modifikationen mit geringer Abundanz wie fdC leichter sequenzieren zu können, ist eine Anreicherung über Antikörper naheliegend. Es gibt sowohl für DNA- als auch für RNA-Modifikationen Sequenzierungsmethoden, welche über Anreicherung die Position der modifizierten Base bestimmen können.^[125, 126, 167] Die aktuellen Antikörper sind jedoch für solche Methoden entweder nicht vorhanden oder besitzen nicht die nötige Spezifität.^[153, 168] Der bisher übliche Weg, um das modifizierte Nukleosid an das gewünschte Trägermolekül zwecks Antikörpergenerierung zu binden, ist die Spaltung des Zuckers mit Periodat gefolgt von einer anschließenden Reduktion der Schiff'schen Base mittels NaBH₄.^[169] Dieser Weg ist für reduktionsempfindliche Basen und die Nukleobasen mit einer 2'-O-Modifikationen nicht geeignet.^[168] Daher sollten in dieser Arbeit monoklonale Antikörper gegen solche Nukleosid-Modifikationen im Oligonukleotid Kontext generiert werden. Monoklonale Antikörper besitzen gegenüber polyklonalen Antikörpern wichtige Vorteile. Zum einen sind monoklonale Antikörper homogen und damit konsistenter als polyklonale Antikörper. Zum anderen ist es möglich monoklonale Antikörper, nach erfolgreicher Etablierung, dauerhaft, in unbegrenzter Menge und immer gleichbleibender Qualität bereit zu stellen. Polyklonale Antikörper hingegen sind je nach Charge von deutlich unterschiedlicher Qualität.^[170] Es sollte eine Kopplungsmethode etabliert werden, die in Kooperation mit der *Monoclonal Antibody Core Facility* des *Helmholtz-Zentrum München*, die Erzeugung monoklonale Antikörper gegen die DNA-Modifikation 5-Formyl-2'-deoxycytidin (fdC) und die RNA-Modifikation 2'-O-Methylguanosin (Gm), ermöglichen sollte. Zu Beginn dieser Arbeit war kommerziell lediglich ein polyklonaler Antikörper gegen fdC (*Active Motif*) erhältlich und kein Antikörper gegen Gm.

Die BS-seq wird nach wie vor als der „Gold Standard“ der einzelbasen-aufgelösten Sequenzierung der epigenetischen Basen bezeichnet. Jedoch kann bei der BS-seq nur mit ssDNA ein quantitativer Umsatz erreicht werden. Die Behandlung von dsDNA liefert nur eine Umsetzung von 0,1% der Cytosine.^[81] Durch die Entwicklung des *hairpin bisulfite sequencing* (HP-BS) von *Laird et al.* konnte dieses Problem geschickt umgangen werden.^[99] *Giehr et. al* weiteten diese Methode kürzlich auf mdC und hmdC aus.^[171, 172] Jedoch fehlt in dieser Methode die Möglichkeit, auch fdC und cadC einzubinden, was für eine umfassende Untersuchung des Methylierungsmusters notwendig wäre. Mit Oligonukleotiden welche alle epigenetischen Basen (mdC, hmdC, fdC und cadC) enthalten sollte es möglich sein die jeweilige spezifische Umwandlungsmethode auf ihre Effektivität hin zu überprüfen. Damit wäre es möglich die Abhängigkeit der dynamische Demethylierung von Hemi- bzw. Homomethylierung, sowohl während der embryonalen Entwicklung als auch in alternden Zellen, besser zu verstehen. In dieser Arbeit sollten daher, *Hairpin*-Oligonukleotide welche die epigenetischen Basen enthalten synthetisiert, charakterisiert und aufgereinigt werden. Anschließend sollten die Oligonukleotide zur Sequenzierung von dsDNA eingesetzt werden.

Das große Interesse an den epigenetischen Basen und deren Sequenzierung bewirkte, dass es beispielsweise für hmdC viele unterschiedliche Methoden zur Sequenzierung mit Bisulfit gibt. Die meisten der Methoden zur Sequenzierung von hmdC haben dabei die Oxidation zu fdC in DNA als Schlüsselschritt gemein. Bei der oxBS wurde für diese Oxidation KRuO_4 verwendet.^[85] Jedoch war schon früh klar, dass es schwierig ist, geeignete Bedingungen zu finden, bei welchen der exzessive Abbau von DNA durch KRuO_4 vermindert sein sollte.^[173, 174] Eine Reproduzierbarkeit der Sequenzierungsdaten ist daher nur schwer zu erreichen. Es gibt verschiedene Ansätze, um den durch KRuO_4 ausgelösten DNA Abbau zu umgehen. Eine prominente Methode ist das *Tet-assisted bisulfite sequencing* (TAB-Seq).^[87] Das TAB-Seq benötigt jedoch hoch aktives TET1 Enzym, welches unter aufwändigen und schonenden Bedingungen isoliert werden muss, damit seine Aktivität erhalten bleibt. Zudem ist bei dieser Methode eine hmdC-Glykosylierung mittels β -Glykosyltransferase als zusätzlicher Schritt notwendig, was zu weiterer Verzerrung der Ergebnisse führen kann. Neben KRuO_4 gibt es noch weitere chemische Oxidationsmittel, für die Oxidation von hmdC in DNA. Eines davon nutzt *2-hydroxy-2-azaadamantane* (AZADOL) und in *sodium dodecyl sulfate* (SDS) Micellen eingeschlossenes *bis(acetoxy) iodobenzene* (BAIB), um die schlechte Wasserlöslichkeit von BAIB zu umgehen.^[175] Eine weitere Methode nutzt eine mittels Tempo/Cu(II) katalysierte Oxidation. Jedoch ist ersichtlich, dass die Oxidation mittels Tempo/Cu(II) in Oligonukleotiden nicht vollständig stattfindet und eine weitere Anreicherung mittels Biotin-markierung notwendig ist.^[176] Zudem wurde diese Methode noch nicht an genomischer DNA getestet. Alle oben genannten Methoden zur Sequenzierung von hmdC haben Limitierungen in ihrer Reaktivität, in der Reproduzierbarkeit der Daten oder der Einsetzbarkeit. Die Oxidation von hmdC zu fdC im Oligonukleotid Kontext sollte daher entscheidend verbessert werden. Um dies zu erreichen, sollten wasserlösliche Derivate des milden und selektiven Oxidationsmittels IBX synthetisiert werden und an Oligonukleotiden auf ihre Effektivität zur selektiven hmdC Oxidation überprüft werden.

Im zweiten Teil dieser Arbeit sollte ich mich mit der Frage beschäftigen, ob ALKBH5 neben m^6A auch das in rRNA natürlich vorkommende m^6_2A , trotz des dimethylierten exozyclischen N^6 -Heteroatoms, demethylieren kann. Die Modifikation m^6_2A befindet sich ausschließlich in der humanen 40S Untereinheit des Ribosoms (18S rRNA; Position 1850 und 1851) in einem hoch konservierten Sequenzkontext.^[121] Dabei sollte diese Studie zeigen, ob ALKBH5 andere Substrate prinzipiell zulässt. Bei einer gegebenenfalls verlangsamten Umsetzung von m^6_2A sollten zudem tiefere Einblicke in den Mechanismus der Demethylierung erzielt werden.

Des Weiteren sollte die Modifikation ms^2A untersucht werden, welche in tRNA vermutet wird. Für den Ursprung von ms^2A in der RNA humaner Zellen, gibt es prinzipiell zwei Möglichkeiten. Zum einen könnte ms^2A direkt mit Hilfe einer Thiomethyl-Transferase als eigenständige Modifikation in die RNA eingebaut werden. Zum anderen könnte ms^2A über eine aktive Dealkylierung von 2-thiomethylierten Adenosin-Modifikationen entstehen. In humanen Zellen beschränken sich die 2-thiomethylierten

Aufgabenstellung

Adenosin-Modifikationen bisher auf $\text{ms}^2\text{i}^6\text{A}$ und $\text{ms}^2\text{t}^6\text{A}$.^[112] Um den Ursprung von ms^2A in humanen Zellen zu identifiziert, sollte im Rahmen dieser Arbeit ein an der Thiomethyl-Gruppe dreifach deuteriertes ms^2A ($\text{d}_3\text{-ms}^2\text{A}$) synthetisiert werden. Das somit markierte Nukleosid sollte in quantitativen LC-MS/MS Studien mit spezifischen *Knockout*-Zelllinien eingesetzt werden.

3. Antikörper gegen RNA- und DNA-Modifikationen

In diesem Teil der Arbeit sollte ein monoklonaler Antikörper gegen die anspruchsvolle Nukleosid-Modifikationen fdC und Gm, im Oligonukleotid Kontext, generiert werden. Der sonst übliche harsche Weg, um das modifizierte Nukleosid an das gewünschte Trägermolekül zu binden, ist für diese Nukleosid-Modifikationen nicht geeignet.^[168]

3.1. Allgemeines Hapten Design

Nukleoside bzw. modifizierte Nukleoside können zwar Antigene sein, aber keine Immunantwort hervorrufen, da sie mit ihrer geringen Größe vom Immunsystem nicht als Antigen erkannt werden oder im Vergleich zu Proteinen zu schnell metabolisiert werden. Diese kleinen Antigene, Haptene, können keine vollwertige Immunantwort hervorrufen. Um aus einem Hapten ein vollwertiges Immunogen zu machen, müssen diese an Makromoleküle immobilisiert werden. Für die spätere EDC-Kopplung an ein Trägermolekül wird am Hapten entweder eine Phosphat- oder eine Carboxylgruppe benötigt. Im Rahmen dieser Arbeit wurde eine Carboxylgruppe als funktionelle Gruppe in das Hapten integriert. Die Generierung von Antikörpern über Phosphat-EDC-Kopplung, direkt an der 5' Position, wurde bereits in der Doktorarbeit von *Veronika Reiter* bearbeitet und scheint besonders bei Hypermodifikationen schwierig zu sein.

Das hier konzipierte Hapten basiert auf zwei Grundgedanken. Zum einen soll die Verbindung von Nukleosid und Carboxylgruppe möglichst mild und zuverlässig sein, um sie auch auf anspruchsvollen Modifikationen anwenden zu können. Zum anderen soll der Abstand der Carboxylgruppe zum Nukleosid variabel sein, damit gesteuert werden kann, wie stark das Hapten auf der Oberfläche des Trägerproteins exponiert ist (Abbildung 15 A und B).

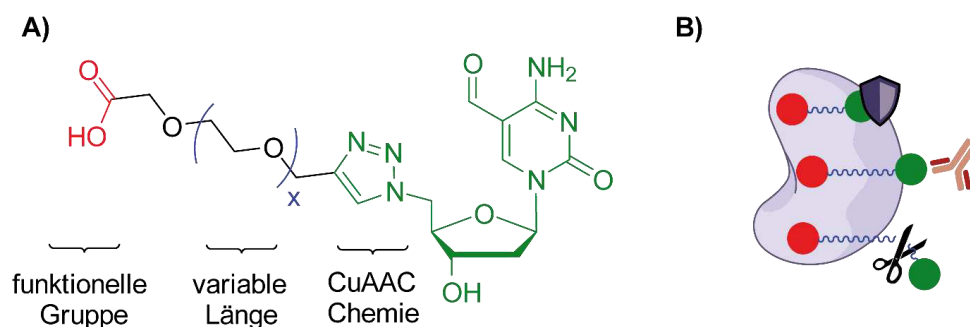


Abbildung 15: A) Übersicht zur Haptenstrategie. Die funktionelle Gruppe ist in rot markiert, der variable Linker in Blau und das Nukleosid mit der Triazol-Verknüpfung (CuAAC) ist Grün dargestellt. B) Schematische Darstellung der Exponiertheit des Haptens gegenüber B-Lymphozyten. Die Funktionelle Gruppe ist in Rot markiert, der variable Linker in Blau und das Nukleosid in Grün dargestellt. Es sind drei Möglichkeiten dargestellt: das Hapten ist vom Trägerprotein verdeckt (lila Schild), das Hapten ist gut exponiert und die B-Lymphozyten können Antikörper generieren (roter Antikörper) und das Hapten ragt aus dem Trägerprotein heraus und wird abgebaut (Schere).

Eine milde und zuverlässige Kopplungsmethode ist die Kupfer(I)-katalysierte Azid-Alkin-Cycloaddition (CuAAC). Durch eine Azidgruppe an der 5'Position des Zuckers des Nukleosids, können verschiedene Linker mit einer Alkingruppe gekoppelt werden. Die CuAAC kann in vielen protischen und aprotischen Lösungsmitteln, wie z.B. Wasser/MeCN Gemischen, durchgeführt werden und ist somit sehr variabel. Ein weiterer großer Vorteil dieser Kopplungsmethode ist ihre Spezifität, damit sind auch bei stark modifizierten Nukleosiden meist keine Schutzgruppen nötig.^[177]

Der variable Abstand zwischen Carboxylgruppe und Nukleosid wird durch Polyethylenglycol (PEG) erreicht werden. PEG erhöht zum einen die Wasserlöslichkeit des Haptens. Zum anderen ist es unter physiologischen Bedingungen stabil und wird beispielsweise in biopharmazeutischen Wirkstoffen benutzt.^[178]

Mit der Länge des *Linkers* kann bestimmt werden, wie weit das Hapten aus der Oberfläche des Trägerproteins herausragen könnte (Abbildung 15 B). Bei zu kleinen Abständen könnte das Nukleosid und damit dessen immunogene Wirkung verdeckt werden. Bei zu großen Abständen könnte das Nukleosid von Enzymen abgespalten werden oder durch den flexiblen PEG-*Linker* verdeckt werden. Für die erste Generation der Hapten-Protein-Konjugate wurde ein Triethylenglykol (PEG3)-*Linker* gewählt. Dies lässt die Möglichkeiten offen, den *Linker* sowohl kürzer (z.B. PEG1) als auch länger (z.B. PEG5) zu gestalten.

3.2. Antikörper der erste Antikörper Generation

Abbildung 16 zeigt die beiden zu synthetisierenden Haptene der ersten Generation.

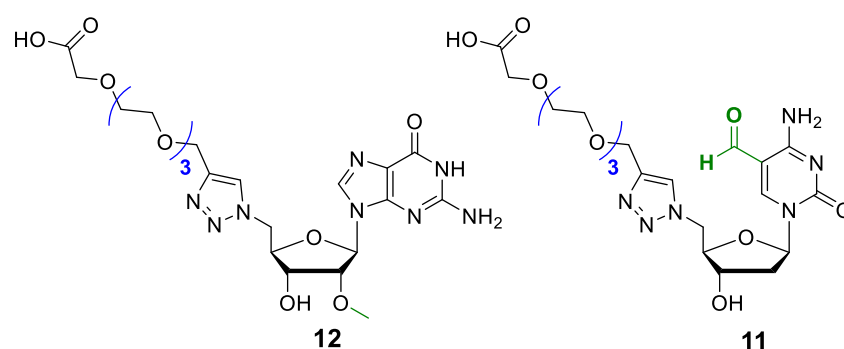


Abbildung 16: Übersicht der in dieser Arbeit synthetisierten Haptene erster Generation: fdC-PEG3-Hapten **11** und Gm-PEG3-Hapten **12**. Die Modifikation ist in Grün und die *Linker* Länge in Blau hervorgehoben.

3.2.1. Synthese der Haptene PEG3-fdC und PEG3-Gm

Zunächst musste für die Synthese der PEG3-fdC und PEG3-Gm Haptene der PEG3-*Linker* **3** dargestellt werden. Die Synthese erfolgte in drei Stufen (Abbildung 17).

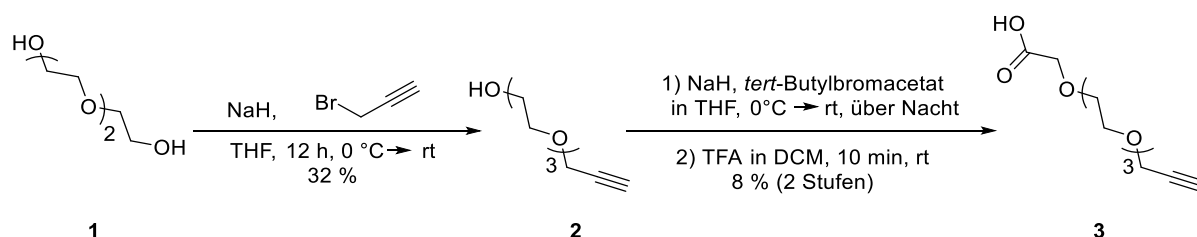


Abbildung 17: Syntheseroute zur Darstellung des PEG3-Linkers **3** ausgehend von Triethylenglykol **1** in 3 Schritten.

Zunächst wurde Triethylenglykol **1** mit Natriumhydrid teilweise deprotoniert und mit Propargylbromid zum monoalkylierten Triethylenglykol **2** umgesetzt (32 % Ausbeute). Um eine Monosubstituierung zu erreichen wurde die Menge der eingesetzten Base (0.65 Äq) sowie die Menge an Propargylbromid (0.5 Äq) reduziert und langsam bei niedriger Temperatur (0 °C) zugetropft. Es wurde jedoch trotz dieser Bemühungen doppelte Substituierung beobachtet. Der freie Alkohol des monoalkylierten Triethylenglykols **2** wurde unter basischen Bedingungen mit *tert*-Butylbromacetat umgesetzt. Nach saurer Entschützung wurde der PEG3-Linker **3** erhalten werden (8 % Ausbeute über zwei Stufen).

Das freie 5-Formylcytidin **4** wurde über die im Arbeitskreis *Carell* etablierte und literaturbekannte CO Insertion dargestellt.^[179, 180] Die Synthese des 5'-Azid-5-formylcytidin **6**, ausgehend vom 5-Formylcytidin **4**, startete mit der Einführung einer *p*-Toluolsulfonsäure Gruppe am 5'-OH des ungeschützten Nukleosids (Abbildung 18).

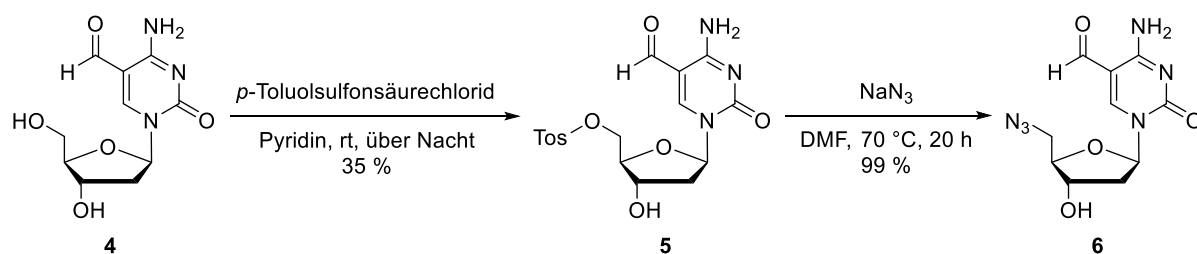


Abbildung 18: Syntheseroute zur Darstellung von 5'-Azid-5-formylcytidin **6** ausgehend von 5-Formylcytidin **4** in 3 Schritten.

So wurde das 5'-Tosly-5-formylcytidin **5** in einer Ausbeute von 35 % als Hauptprodukt erhalten. Die Tosylgruppe als gute Abgangsgruppe konnte im weiteren Schritt mit Natriumazid quantitativ zu 5'-Azid-5-formylcytidin **6** umgesetzt werden.

Die Synthese des Phenoxyacetyl (PAC)-geschützten 5'-Azid-2'-*O*-methylguanosin **10** wurde vom kommerziell erhältlichen 2'-*O*-Methylguanosin **7** gestartet (Abbildung 19).

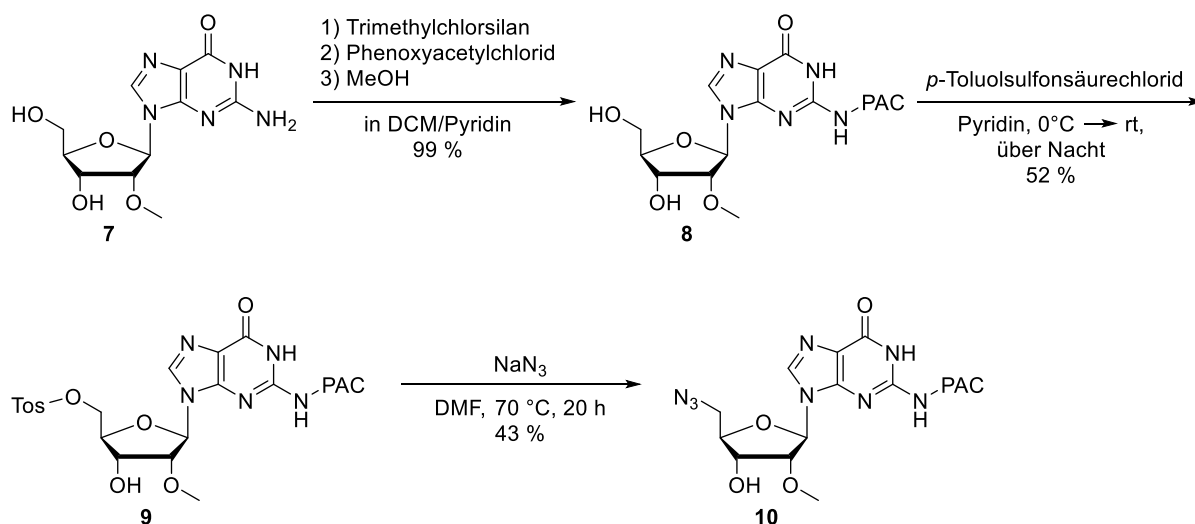


Abbildung 19: Syntheseroute zur Darstellung von 5'-Azid-2'-*O*-methylguanin **10** mit PAC-Schutzgruppe, ausgehend vom 2'-*O*-Methylguanin **7** in drei Schritten.

Um die Löslichkeit zu erhöhen wurde das 2'-*O*-Methylguanin im ersten Schritt am exozyklischen Amin mit PAC geschützt. Angelehnt an *Fan et al.* [181] wurde zuerst das gesamte Nukleosid transient silylgeschützt. Erst im Anschluss erfolgte die Umsetzung mit PAC-Chlorid. Damit konnte 2N-PAC-2'-*O*-methylguanin **8** in quantitativer Ausbeute erhalten werden. Analog zum fdC-Derivat konnte die Einführung einer *p*-Toluolsulfonsäure-Gruppe am 5'-OH des ungeschützten Zuckers erreicht werden (**9**, 52 % Ausbeute). Das Tosylat konnte im weiteren Schritt, mit Natriumazid, zu **10** mit 43 % Ausbeute umgesetzt werden. Ein Temperatur und Reaktionszeit-Screening ergab keine Verbesserung und zeigte teilweise Zersetzung bei höheren Temperaturen.

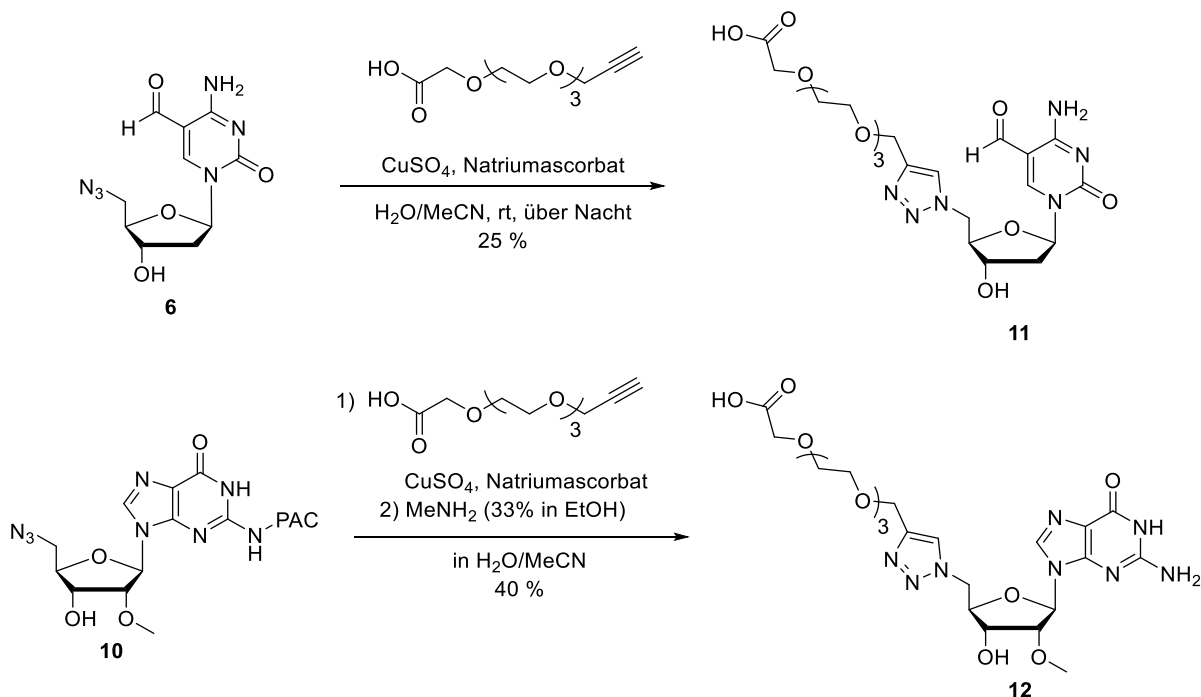


Abbildung 20: Darstellung von fdC-PEG3-Hapten **11** und Gm-PEG3-Hapten **12** ausgehend von den 5'-Azid-Nukleosiden **6** bzw. **10** mittels CuAAC mit dem PEG3-Linker **3**.

Über eine CuAAC zwischen fdC-Azid **6** und dem Alkin des PEG3-*Linker* **3** wurde das fdC-PEG3-Hapten **11** (25 % Ausbeute) erhalten (Abbildung 20). Das Gm-PEG3-Hapten **12** wurde analog dazu aus dem Gm-Azid **10** erhalten. Die PAC-Schutzgruppe des Gm-Haptens wurde anschließend vollständig mit MeNH₂ entfernt. Es wurde dabei keine Transaminierung beobachtet und nur das Hydrolyse-Produkt erhalten (40 % Ausbeute). Beide Haptene wurden mittels RP-HPLC zweifach gereinigt.

3.2.2. EDC-Kopplung der Haptene **11** und **12** mit OVA

Die gängigste Methode zur Kopplung von Nukleosiden an Trägerproteine ist die oxidative Periodatspaltung mit anschließender Reduktion.^[169] Dabei wird der Zucker des Nukleosids gespalten und die daraus resultierenden Aldehyd-Gruppen mit Lysinresten des Trägerproteins umgesetzt und mit Hilfe von NaBH₄ reduziert, um die Verbindung zu stabilisieren (siehe Abbildung 21).

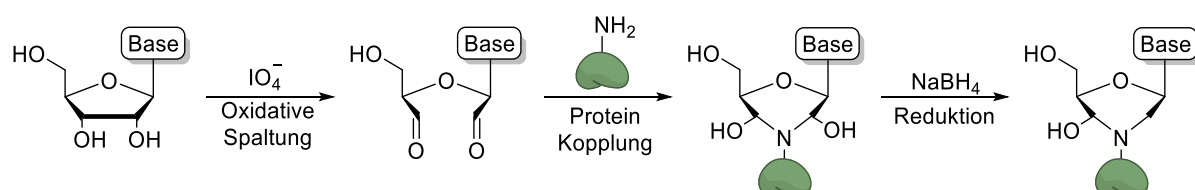


Abbildung 21: Kopplung mittels Periodatspaltung. Im ersten Schritt der Kopplung wird der Zucker des Nukleosids am Diol (2'-OH und 3'-OH) oxidativ gespalten. Die entstandenen Aldehyde können mit Lysin-Resten des Trägerproteins reagieren. Über eine Reduktion mit NaBH₄ wird die Bindung zu einer stabilen kovalenten Verknüpfung.^[169]

Bei den in dieser Arbeit adressierten modifizierten Nukleosiden ist diese Methode jedoch keine Option. Gm ist nicht für die Periodatspaltung geeignet. Es ist an der 2'-OH-Gruppe des Zuckers methyliert und bietet somit kein freies Diol. Darüber hinaus, würde eine Spaltung des Zuckers das Hapten zerstören. Die Formyl-Gruppe des fdC hingegen ist voraussichtlich bereits bei den harschen Reaktionsbedingungen der Periodat-Oxidation nicht stabil. Da für die spätere Reduktion NaBH₄ verwendet wird, käme es spätestens dann zur Bildung von hmdC aus fdC.^[88] Zudem könnte die Formyl-Gruppe des fdC mit den Lysinresten des Trägerproteins reagieren.

Eine sehr viel mildere und ebenso etablierte Methode ist die Kopplung des Nukleosids mittels 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimid (EDC). Das EDC reagiert mit der Carboxylgruppe zu einem aktivierten Isoharnstoff-Intermediat. Das Isoharnstoff-Intermediat kann im Folgenden mit einem Aminrest des Trägerproteins zum stabilen Amid umgesetzt werden (Abbildung 22). Diese Reaktion ist sehr robust und kann in bis zu 20 % DMSO in Wasser durchgeführt werden.^[182, 183, 184, 185]

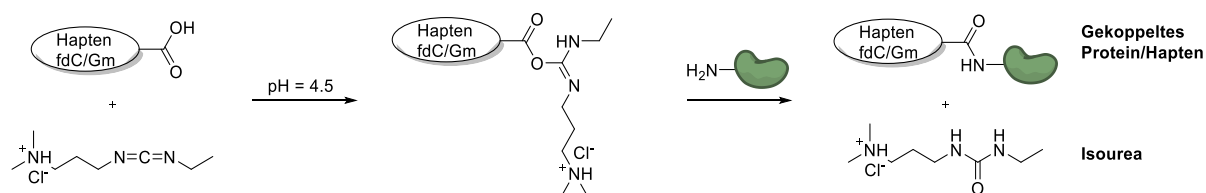


Abbildung 22: Mechanismus der EDC-Kopplung. Die Säuregruppe des Haptens wird zunächst mit EDC aktiviert. Durch anschließenden nukleophilen Angriff einer Aminogruppe des Trägerproteins entsteht eine kovalente Peptidbindung.

fdC-PEG3-Hapten **11** und Gm-PEG3-Hapten **12** wurden daher mittels EDC an das Trägerprotein OVA gekoppelt. OVA ist der Hauptbestandteil von Eiklar (Vogel), hat eine Molekülmasse von 45 kDa und präsentiert 20 Lysine an der Oberfläche.^[186, 187] Für die erste Generation der Hapten-Protein-Konjugate wurde OVA genutzt, da das *Monoclonal Antibody Core Facility* des *Helmholtz-Zentrum München* mit diesem Protein bisher die besten Ergebnisse erreicht hat. Ein großer Vorteil von OVA ist zudem die hohe Löslichkeit in DMSO (bis zu 70 %).^[185] Damit lassen sich auch Haptene mit OVA koppeln, die sich schlecht in Wasser, aber gut in Wasser/DMSO-Gemischen lösen. Im Hinblick auf die besonders geringe Löslichkeit von Guanodin-Derivaten in Wasser, ist diese Eigenschaft sehr hilfreich.

Das kommerziell erhältliche Imject® OVA von *Thermo Fisher* wurde dafür mit einem MES-Puffer auf den pH-Wert 4.6 eingestellt. Im sauren Milieu aktiviert EDC den Säurerest des Haptens und bildet einen aktivierten Ester (*O*-Acylisourea). Damit kann ein Nukleophil, wie die freie Aminogruppe einer Lysin-Seitenkette des Trägerproteins, mit dem aktivierten Ester zu einem stabilen Peptid reagieren (Abbildung 22).

Die Kopplung des fdC-PEG3-Hapten **11** verlief problemlos und das fdC(PEG3)-OVA-Konjugat konnte mittels einer Entsalzungs-Säule (*Zeba™ spin desalting column*) von überschüssigem Hapten getrennt werden. Die Proteinkopplung mit dem Gm-PEG3-Hapten **12** war durch die schlechte Löslichkeit in Wasser eine Herausforderung. Mit der Zugabe von 20% DMSO, konnte **12** zwar gelöst werden, aber während der Kopplung bildete sich ein Niederschlag. Dieser konnte nicht wieder in Lösung gebracht werden. Da DMSO bei den weiteren Immunisierungsexperimenten stören kann, wurde ein Pufferaustausch mittels Entsalzungs-Säulen durchgeführt. Trotz des harschen Pufferwechsels konnte das Gm-PEG3-OVA-Konjugat in Lösung erhalten werden. Die Ausbeute lag bei noch 50 % des eingesetzten Proteins.

3.2.3. Charakterisierung der EDC Kopplungen der ersten Generation

Die erfolgreiche Bildung von Hapten-OVA-Konjugaten wurde mithilfe von *Michael Stadlmaier* mittels Massenspektrometrie und UV/Vis qualitativ untersucht. Die Hapten-OVA-Konjugate wurden mit Trypsin verdaut und die entstandenen Peptide auf Modifizierungen untersucht. Die Daten wurden mit *Proteom Discoverer 2.2* ausgewertet und strikt mit einer *false discovery rate* (FDR) von 1 % gefiltert. Die Abbildungen 23 und 24 zeigen die Aminosäure-Sequenz von OVA und die identifizierten Hapten-Peptid-Konjugate (grün hinterlegt).

	1	11	21	31	41	51	61	71	81	91
1	M			C	O	M				M
	M									
101			C M						M	O
201	M	O						M M		
301										

Abbildung 23: Übersicht der mit *Proteom Discoverer 2.2* identifizierten Hapten-Peptid-Konjugate der Kopplung zwischen fdC-PEG3-Hapten **11** und OVA (grün hinterlegt). Die verwendete *false discovery rate* (FDR) lag bei 1 % und die *coverage* betrug 28.5 %. M = fdC-PEG3 Modifikation an einem Lysin; C = Carbamidomethyl; O = Oxidation von Methionin. Die Modifikationen C und O wurden nur aufgeführt, wenn sie in einer mit M modifizierten Peptidsequenz auftraten.

	1	11	21	31	41	51	61	71	81	91
1	M									M
	M									
101			C M						M	O
201	M	O						M		
301										

Abbildung 24: Übersicht der mit *Proteom Discoverer 2.2* identifizierten Hapten-Peptid-Konjugate der Kopplung zwischen Gm-PEG3-Hapten **12** und OVA (grün hinterlegt). Die verwendete *false discovery rate* (FDR) lag bei 1 % und die *coverage* betrug 19.4 %. M = Gm-PEG3 Modifikation an einem Lysin; C = Carbamidomethyl; O = Oxidation von Methionin. Die Modifikationen C und O wurden nur aufgeführt, wenn sie in einer mit M modifizierten Peptidsequenz auftraten.

Die vergleichende UV/Vis Messungen zwischen OVA und dem Hapten-OVA-Konjugat, zeigten einen Anstieg der Absorption bei 260 nm, dem Absorptionsmaximum für Nukleoside.

3.2.4. Optimierung des ELISA-Screenings mit biotinylierten Oligonukleotiden

Nach der Immunisierung von Ratten mit den Hapten-OVA-Konjugaten (fdC und Gm), wurden Hybridomzellen generiert. Die im Medium-Überstand enthaltenen sekretierten Antikörper wurden über ein ELISA-Screening selektiert. Für das ELISA-Screening, werden üblicherweise Bruchstücke der mit Hapten-konjugierten Proteine verwendet, mit Biotin gekoppelt und auf Streptavidin-Platten immobilisiert. Mit diesen wird dann nach potenziellen Antikörpern gesucht.^[188] Das Ziel der hier synthetisierten Antikörper waren aber modifizierte Nukleoside im DNA bzw. RNA Kontext. Daher wurden in dieser Arbeit Oligonukleotide mit Biotin-Tag für das ELISA-Screening entworfen, um diese Vorauswahl zu verbessern.

Für die RNA-Modifikation Gm wurde ein Kompromiss eingegangen. RNA kann, vor allem in Anwesenheit von Metallionen, hydrolysiert werden.^[189] Das Medium der Hybridomzellen (im Weiteren Hybridomzellen-Überstand) beinhaltet zudem stets auch RNasen, welche zum Abbau eines RNA-Antigens führen würden. Um den Abbau zu verhindern wurde die Gm Modifikation in DNA eingebettet und somit ein RNA/DNA-Hybrid hergestellt.

3.2.5. Festphasensynthese der Oligonukleotide für das ELISA-Screening

Die genutzten DNA bzw. RNA/DNA-Hybride mit Biotin-Tag wurden mittels Festphasensynthese synthetisiert. Die Sequenzkontexte der beiden Modifikationen sind sehr variabel. Daher wurden die umliegenden DNA-Basen randomisiert eingebaut, um möglichst alle statistisch möglichen Sequenzen abzudecken. Dies wurde erreicht, indem die vier kanonischen Nukleosid-Phosphoramidite äquimolar vorgemischt wurden. Dieses Gemisch wurde dann im Standardprotokoll der Festphasensynthese genutzt. Die Sequenzen wurden bewusst mit 6 bzw. 8 Basen entworfen (siehe Tabelle 1), da längere Oligonukleotide anfälliger für den Abbau durch RNasen bzw. DNasen sind. Auch ist die Aufreinigung und Charakterisierung von kürzeren randomisierten Oligonukleotiden leichter, was eine höhere Ausbeute zur Folge hat.

Tabelle 1: Sequenzen der für das ELISA-Screening synthetisierten randomisierten Oligodesoxyribonukleotide (ODN). Modifizierte Basen sind fett markiert. N = dA/dC/dG/T.

ODN #	Sequenz
ODN 1	5'-Biotin N fd CN fd CNC-3'
ODN 2	5'-Biotin NCN CNC-3'
ODN 3	5'-Biotin NNN fd CNN NA-3'
ODN 4	5'-Biotin NNN CNN NA-3'
ODN 5	5'-Biotin NNN Gm NN NC-3'
ODN 6	5'-Biotin NNN rGNN NC-3'

Für die Modifikation Gm wurde daher das Motiv 5'-Biotin NNN GmNN N-3' und das unmodifizierte Motiv 5'-Biotin NNN rGNN N-3' als Negativkontrolle gewählt (N = dA, dC, dG und T). Dieses Sequenzmotiv bestehend auch 7 Nukleobasen, sollte genügend Raum für die Bindung des potentiellen Antikörpers bieten. Für die Modifikation fdC wurden entsprechend die Motive 5'-Biotin NNN fdCNN NA-3' und 5'-Biotin NNN CNN NA-3' gewählt. Zudem wurden die Sequenzen 5'-Biotin NfdCN fdCNC-3' und 5'-Biotin NCN CNC-3' synthetisiert, um eine höhere Modifikationsdichte zu erreichen.

Die ODNs **3,4,5** und **6** enthalten jeweils 6 randomisierte Positionen. Damit ergeben sich 4096 mögliche individuelle Sequenzen. Aufgrund der unterschiedlichen Basenzusammensetzung und der Ausbildung von Sekundärstrukturen, weisen die ODNs bei der Auftrennung mittels RP-HPLC keine klar definierten Signale auf. Ähnliches gilt für die ODNs **1** und **2**, welche aus 64 individuelle Sequenzen bestehen. Daher wurde der Biotin-Tag über das Phosphoramidit **13** (Abbildung 25), als letzter Baustein in der 5'-Position eingefügt.

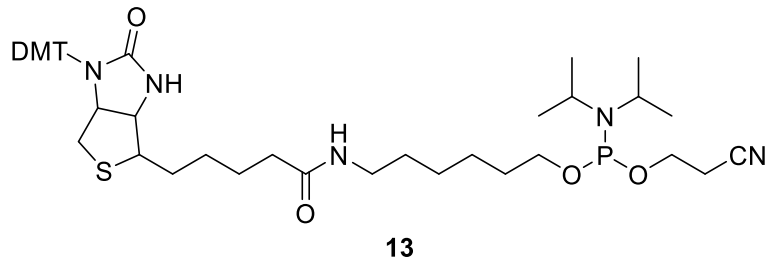
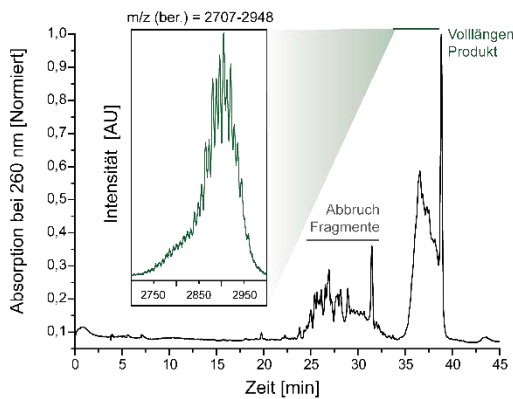


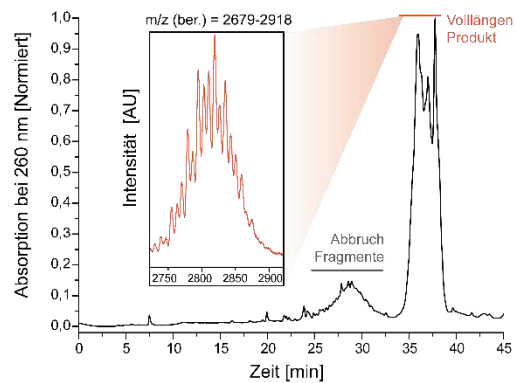
Abbildung 25: Das Biotin Phosphoramidit **13**, welches in der Festphasensynthese genutzt wurde. Mit diesem wurde der Biotin-Tag an der 5'-Position der Oligonukleotide eingefügt.

Der Einbau an letzter Position hat vor allem Vorteile in der Aufreinigung mittels RP-HPLC. Biotin verursacht hier eine starke Verschiebung der Retention zu späteren Elutionszeiten. Ähnlich zu DMT-*on* Synthesen lässt sich so das Volllängenprodukt von den Abbruchfragmenten trennen, die bei der Festphasensynthese anfallen.

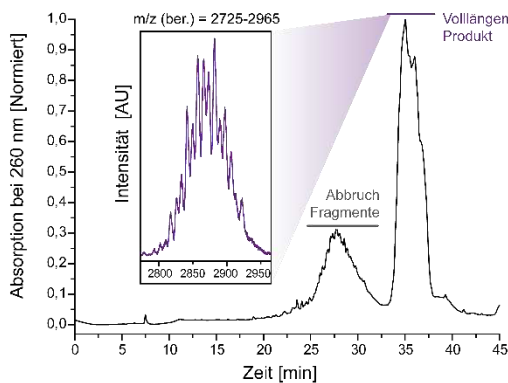
A) fdC



B) dC



C) Gm



D) G

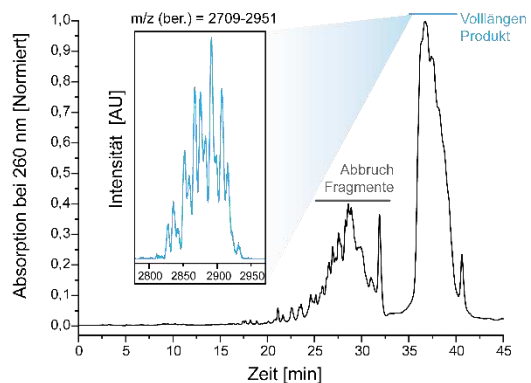


Abbildung 26: Charakterisierung der randomisierten Oligonukleotide mittels HPLC und MALDI. Die ODNs vor der Aufreinigung sind schwarz dargestellt. Die Massenspektren der reinen Fraktionen sind jeweils schwarz umrandet und farbig gekennzeichnet (grün = fdC; orange = dC; lila = Gm; blau = G). Dargestellt sind die semipräparativen RP-HPLC Chromatogramme der DNA-Oligonukleotide: **A)** 5'-Biotin NNN **fdC**NN NA-3' (grün) und **B)** 5'-Biotin NNN **dC**NN NA-3' (orange). Sowie die semipräparativen RP-HPLC Chromatogramme der DNA/RNA-Hybride: **C)** 5'-Biotin NNN **rGm**NN NC-3' (lila) und **D)** 5'-Biotin NNN **rG**NN NC-3' (blau).

In Abbildung 26 ist der, durch Biotin verursachte, deutliche *shift* klar zu erkennen. Die Verschiebung der Elutionszeit der Volllängenprodukte betrug bis zu 10 min, gegenüber der Abbruchfragmente. Damit ließen sich die Abbruchfragmente zuverlässig abtrennen. Die Charakterisierung mittels MALDI-Massenspektrometrie zeigte beim dC-ODN **4** (Abbildung 26 B) eine Massenverteilung zwischen der Masse des Oligonukleotids mit der geringsten Masse (5'-Biotin CCC CCC CA-3', 2679 m/z) und der Masse des schwerst möglichen Oligonukleotids (5'-Biotin GGG CGG GA-3', 2918 m/z). Die Signale der geringsten sowie höchsten Massen sind dabei nicht besonders ausgeprägt, da diese an den Rändern der Verteilung liegen und damit stochastisch betrachtet seltener auftreten. Auch das fdC ODN **3** zeigte eine erwartete Massenverteilung zwischen den Extremwerten 2707 m/z (nur C) und 2948 m/z (nur G) (Abbildung 26 A). Die Synthese des Gm ODN **5** und G ODN **6** zeigten relativ viele Abbruchfragmente (Abbildung 26 C und D), welche aber wie schon vorher beschrieben gut abzutrennen waren. Die Massenbereiche lagen auch hier im erwarteten Bereich.

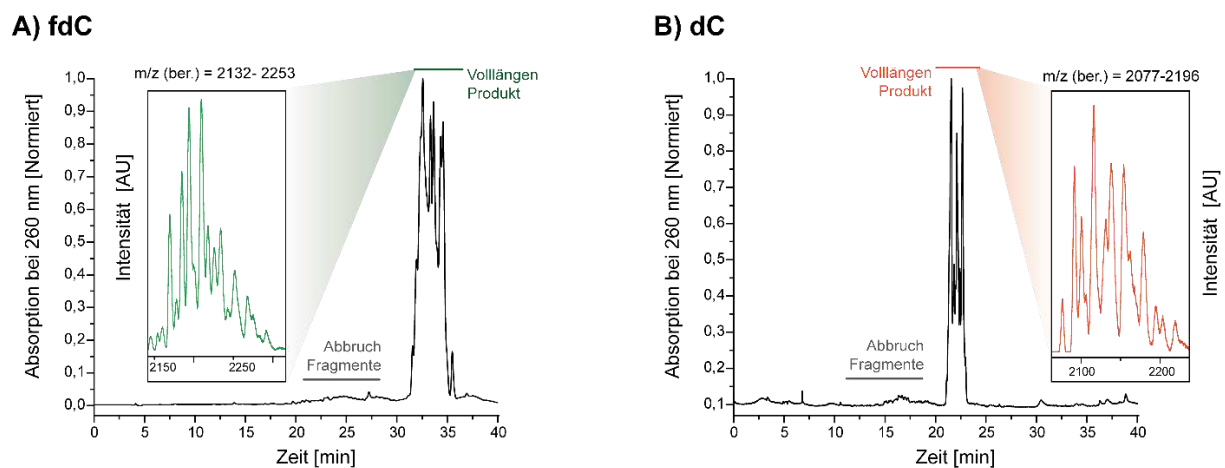


Abbildung 27: Charakterisierung der randomisierten Oligonukleotide mittels RP-HPLC und MALDI. Die ODNs vor der Aufreinigung sind schwarz dargestellt. Die Massenspektren der reinen Fraktionen sind jeweils schwarz umrandet und farbig gekennzeichnet (grün = fdC; orange = dC). Dargestellt sind die semipräparativen RP-HPLC Chromatogramme der DNA-Oligonukleotide: A) 5'-Biotin NfdCN fdCNdC-3' (grün) und B) 5'-Biotin NCN CNC-3' (orange).

Bei der Synthese vom fdC ODN **1** und ODN **2** zeigten sich interessanterweise kaum Abbrüche. Ein Grund dafür könnte sein, dass diese Sequenz nur drei randomisierte Positionen besitzt. Zudem ist die Sequenz kürzer. Damit werden weniger sekundäre Strukturen möglich, die zu vermehrten Abbrüchen führen können. Die Charakterisierung mittels MALDI-Massenspektrometrie zeigt eine erwartete Verteilung zwischen 2132 m/z und 2253 m/z für das fdC ODN **1** und zwischen 2077 m/z und 2196 m/z für das dC ODN **2** (Abbildung 27 A und B).

Beide Oligonukleotid-Varianten von fdC bzw. die Variante von Gm (siehe Tabelle 1) konnten erfolgreich synthetisiert werden und im ELISA-Screening verwendet werden. Die Oligonukleotide wurden nach der Synthese entsalzt und lyophilisiert bei -20 °C aufbewahrt.

3.2.6. Optimierung der Dot Blot Bedingungen

Die im ELISA-Screening gegen die jeweilige Modifikation positiv getesteten Hybridomzellen-Überstände, wurden mittels Dot Blot auf ihre Spezifität (bezüglich der weiteren epigenetischen Basen) getestet. Dabei wurden synthetisch hergestellte Oligonukleotide mit der Zielmodifikation auf eine kationische, verstärkte Nylon-Membran gebunden und mit dem Hybridomzellen-Überstand inkubiert. Die Nylon-Membran weist im Vergleich zu Nitrozellulose eine höhere Stabilität und Reißfestigkeit auf. Die positiven Ladungen der Membran unterstützt dabei die Immobilisierung von Oligonukleotiden durch polare Bindungen mit dem teils negativ geladenen DNA-Rückgrat. Die immobilisierten Oligonukleotide enthielten die Zielmodifikation oder eine der weiteren epigenetischen Modifikationen. Der primäre Antikörper aus den Hybridomzellen-Überständen sollte die Zielmodifikation binden. Dieser Antikörper aus den Hybridomzellen-Überständen wurde anschließend mit einem sekundären subtypenspezifischen anti-IgG Antikörper markiert. Der sekundäre Antikörper konnte in Verbindung mit einem Chemilumineszenz-Reagenz sichtbar gemacht werden und gab quantitative Auskunft über das Bindungsverhalten.

3.2.6.1. Auswahl der Oligonukleotide

Um eine möglichst konstante Auswertung zu ermöglichen, wurden Vorversuche mit dem kommerziell erhältlichen polyklonalen fdC-Antikörper und mit der ersten Charge an Hybridomzellen-Überstände gemacht. Im ersten Schritt wurde hierbei die Länge der DNA Oligonukleotide untersucht. Es standen drei verschiedene Oligonukleotide zur Verfügung, welche von *Edris Parsa* zur Verfügung gestellt wurden:

A = 5'-GTA ATG XGG TAT GG-3' (X = dC, mdC oder fdC)

B = 5'-ATA ATA XGA CTC ACT ATA GGG-3' (X = dC, mdC oder fdC)

C = 5'-TAG TCA CTG CTC GGG ATT CCC TAT AGT GAG TYG TAT TAT-3' (Y = dC, mdC, hmdC, fdC oder cadC)

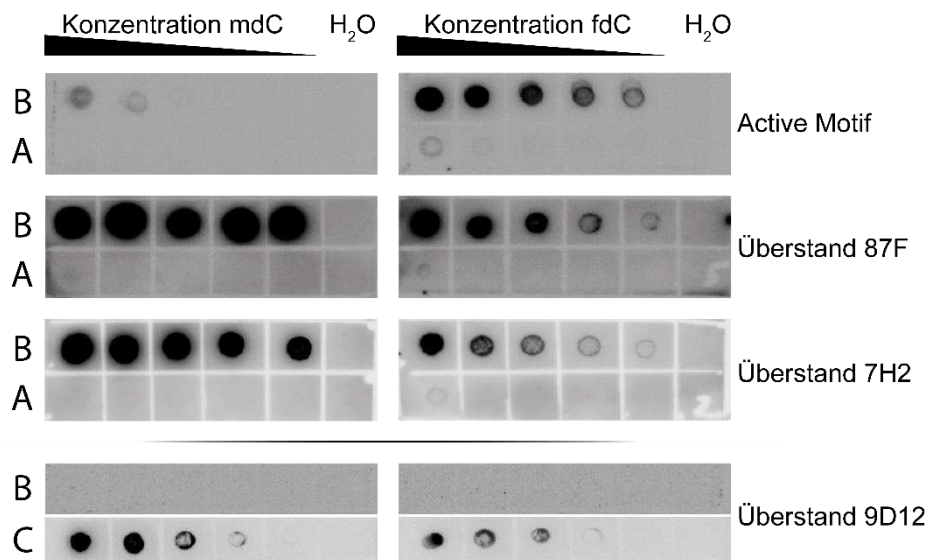


Abbildung 28: Vergleich der Signalintensität, mit den zur Verfügung stehenden Oligonukleotid Sequenzen A, B und C. Auf die Membranen der linken Seite wurden die Oligonukleotid Sequenzen (A, B oder C) mit der Modifikation mdC aufgetragen. Auf der rechten Seite wurden die Oligonukleotid Sequenzen (A, B oder C) mit der Modifikation fdC aufgetragen. Ausgehend von $c = 21 \mu\text{M}$ (erste Spalte) wurden diese jeweils um die Hälfte verdünnt aufgetragen. Nebenstehend sind der verwendete Antikörper/Hybridomzellen-Überstand aufgeführt. Die Oligonukleotide wurden für 1 h bei 80°C auf der Nylon-Membran fixiert.

In den ersten Dot Blots wurde schnell klar, dass eine längere Sequenz eine bessere Bindung an die Membran bedeutet (Abbildung 28). Es konnte bei dem Überstand **87F** und **7H2** und bei dem kommerziell erhältlichen fdC Antikörper (*Active Motif*) mit der kurzen Sequenz A keine Bindung beobachtet werden. Mit der längeren Sequenz B zeigte sich aber eine starke Bindung, sowohl für mdC als auch für fdC. Die längste Sequenz C (39 Basen) stand erst später zur Verfügung und wurde daher vergleichend nur an Überstand **9D12** getestet, da die Überstände der vorherigen Charge aufgebraucht waren. Hier ist nochmal eine klare Verbesserung der allgemeinen Bindung zwischen Sequenz B (keine sichtbare Bindung) und Sequenz C (sichtbare Bindung) zu erkennen. Die darauf folgenden Experimente zur Antikörpervalidierung wurden daher alle mit der längsten Sequenzvariante C durchgeführt.

Tabelle 2: Die in dieser Arbeit für die Dot Blots genutzten 2'-O-Methyl-modifizierten Oligonukleotide. Die Oligonukleotide wurden von der Firma *Ella Biotech* bezogen.

ORN #	Sequenz
ORN EB-1	5'-CUC AGU GCU Gm GC AUA GAU CGA-3'
ORN EB-2	5'-CUC AGU GCU GCm C AUA GAU CGA-3'
ORN EB-3	5'-CUC AGU GCU GCAm AUA GAU CGA-3'
ORN EB-4	5'-CUC AGU GCU GCA Um UA GAU CGA-3'

Für die Dot Blots der Gm-Hybridomzellen-Überstände, standen die in Tabelle 2 gelisteten Oligonukleotide ORN **EB-1** – **EB-4** zur Verfügung. Es wurde darauf geachtet, dass die verwendeten

Oligonukleotide keine Sekundärstrukturen ausbilden sollten. Jedoch stand hier kein kommerzieller Antikörper für Vergleichstest zur Verfügung.

3.2.6.2. Auswahl der Crosslink-Methode

Eine weitere Variable bei Dot Blots bietet das *crosslinking* der Oligonukleotide mit der Nylonmembran. Eine Methode, um Oligonukleotide an Nylon-Membranen zu koppeln, ist das *crosslinking* mit thermischer Energie. Die Membran wurde dabei für 1 h bei 80°C gebacken und das Oligonukleotid an diese gebunden.^[190, 191] Eine weitere Methode ist das *crosslinking* mit UV Strahlung, dabei wurde die Nylon Membran mit UV Licht für 10 min bestrahlt.^[192, 193] Das UV Licht aktiviert in der DNA und RNA die Basen Thymin bzw. Uracil und diese bilden eine kovalente Bindung zur Membranoberfläche aus. Durch diese Bindung wird das Signal der Oligonukleotide, im Vergleich zur Verknüpfung mittels thermische Energie, stark erhöht.^[192, 194]

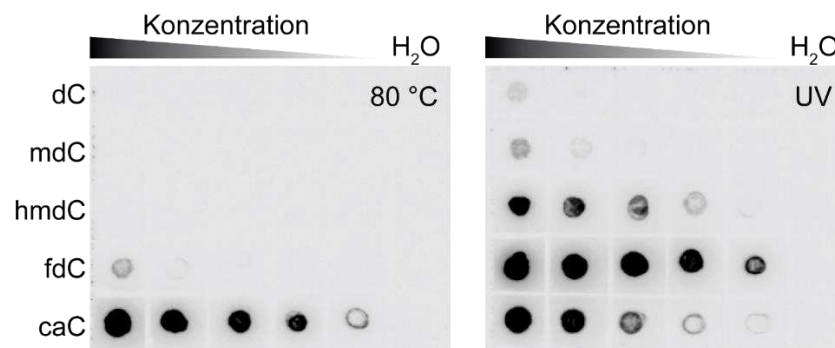


Abbildung 29: Vergleich der Effektivität der Membranfixierung, mittels kommerziell erhältlichen fdC-Antikörper (*Active Motif*). Auf der linken Seite wurde die Nylon-Membran für 1 h bei 80 °C inkubiert. Auf der rechten Seite wurde die Nylon Membran für 10 min mit UV-Licht bestrahlt. Die Oligonukleotide mit den unterschiedlichen Modifikationen (dC, mdC, hmdC, fdC und cadC) wurden ausgehend von $c = 21 \mu\text{M}$ (erste Spalte) jeweils um die Hälfte verdünnt aufgetragen. In der letzten Spalte wurde als Negativkontrolle ddH₂O aufgetragen.

Wie aus Abbildung 29 ersichtlich, liefert die Fixierung über UV Strahlung, vor allem im Bereich mdC, hmdC und fdC, sehr viel bessere Ergebnisse. Daher wurden im Weiteren die Oligonukleotide ausschließlich über UV Bestrahlung auf der Nylonmembran fixiert.

3.2.7. Auswertung der Dot Blots - Überstände erste Generation

Tabelle 3: Übersicht der in Dot Blots getesteten Hybridomzellen-Überstände der ersten Antikörper Generation. Die Selektivität in den Dot Blots wurde durch eine Verdünnungsreihe eingeordnet in schwach, mittel und stark. DNA-E steht für die Modifikation fdC (fett markiert) und DNA-F steht für Modifikation Gm (fett markiert).

Bezeichnung	IgG Subklasse (Ratte)	Selektivitäten (im Dot Blot)
DNA-E R2C 5fC 2C4	IgG2c	mdC (mittel)
DNA-E R2C 5fC 7H2 + BB	IgG2c	mdC (stark), fdC (mittel), dC (schwach)
DNA-E R2C 5fC 1D3	IgG2c	mdC (stark)
DNA-E R2C 5fC 18C10	IgG2c	mdC (mittel)

DNA-E R2C 5fC 8F7 + BB	IgG2c	mdC (stark), fdC (mittel), dC (schwach)
DNA-E R2A 5fC 7H4	IgG2a	Keine Signale
DNA-E R2B 5fC 3C9	IgG2b	Keine Signale
DNA-E R2A 5fC 4H9	IgG2a	Keine Signale
DNA-E R2A 5fC 9B9	IgG2a	Keine Signale
DNA-E R2A 5fC 3F8	IgG2a	Keine Signale
DNA-E R2A 5fC 9H9	IgG2a	Keine Signale
DNA-E R2A 5fC 9D12	IgG2a	dC, mdC, hmdC, fdC, cadC (mittel)
DNA-F R2b mG 9C12	IgG2b	Keine Signale
DNA-F R2b mG 7A6	IgG2b	Keine Signale
DNA-F R2b mG 4A10	IgG2b	Keine Signale
DNA-F R2b mG 17H7	IgG2b	Keine Signale

Die Ergebnisse der einzelnen Dot Blots der ersten Generation wurden in Tabelle 3 zusammengefasst. Für die Hybridomaüberstände **7H2+BB** und **8F7+BB** konnte eine Bindung an fdC gezeigt werden. Jedoch war die Bindung an mdC um ein Vielfaches stärker. Zudem zeigten beide Überstände eine Bindung an dC. Daher sind sie als Antikörper gegen fdC ungeeignet. **2C4**, **1D3** und **18C10** zeigten in den Dot Blots ausschließlich eine Bindung gegenüber mdC. Keine Bindung zeigten **7H4**, **3C9**, **4H9**, **9B9**, **3F8** und **9H9**.

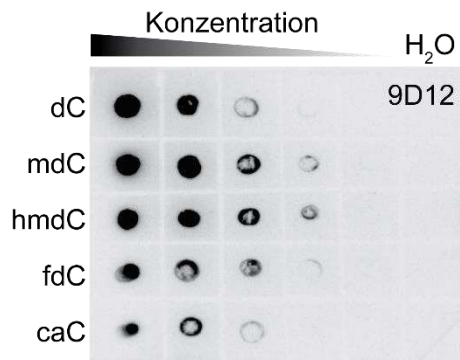


Abbildung 30: Dot Blot des Hybridomaüberstands 9D12. Die Oligonukleotide mit den unterschiedlichen Modifikationen (dC, mdC, hmdC, fdC und cadC) wurden ausgehend von 21 µM (erste Spalte) jeweils um die Hälfte verdünnt aufgetragen. In der letzten Spalte wurde ddH₂O als Negativkontrolle aufgetragen.

9D12 zeigte keine Spezifität und konnte das unmodifizierte und alle modifizierten Oligonukleotide in ähnlicher Stärke binden. Eine leichte Tendenz der stärkeren hmdC Bindung ist dabei zu erkennen. Vermutlich erkennt dieser Antikörper die DNA allgemein und macht dabei weniger basenspezifische Unterschiede. (Abbildung 30)

Wie aus Tabelle 3 ersichtlich, wurden mit den Hybridomzellen-Überständen mit mutmaßlichen Gm-Antikörpern **9C12**, **7A6**, **4A10** und **17H7** im Dot Blot keine Bindung an die verwendeten Oligonukleotide erreicht. Interessanterweise konnten im ELISA-*Screening* nur diese vier Überstände, als Kandidaten identifiziert werden. Möglicherweise war das Konjugat aus Gm-Hapten und OVA nur schwach immunogen. Auch eine Verlängerung der Sequenzen im ELISA-*Screening* und in den Dot Blots könnte sich, ausgehend von den Erfahrungen mit fdC, positiv auswirken.

3.3. Antikörper der zweiten Generation

Aus der ersten Generation konnten viele wertvolle Ergebnisse gesammelt werden, um die Methode weiter entwickeln zu können. Im folgenden Abschnitt soll die Herstellung von monoklonalen Antikörpern mit Hilfe dieser Erkenntnisse optimiert werden.

In der ersten Generation der Antikörper konnte beobachtet werden, dass die Immunantwort relativ schwach ausgefallen ist. Daher sollte ein weiteres Hapten synthetisiert werden. Zudem sollte mit kationisierte bovine Serumalbumin (cBSA) ein Trägerprotein zur Konjugation eingesetzt werden, welches eine stärkere Immunantwort hervorrufen sollte. Es konnten in der ersten Generation bereits Antikörper gegen fdC generiert werden, die jedoch auch starke Kreuzreaktivitäten gegen mdC, hmdC und cadC zeigen. Daher sollten auch die im ELISA-Screening eingesetzten Oligonukleotide verbessert werden.

Die Optimierung der Methode wurde zudem auf fdC als Antigen beschränkt. Für fdC ist bereits ein polyklonaler Antikörper bekannt, damit kann davon ausgegangen werden, dass fdC als Antigen geeignet ist und eine ideale Plattform bietet, um unsere Methode weiterzuentwickeln. Deshalb wurde auf die Weiterentwicklung des Gm Haptens vorerst verzichtet (Abbildung 31).

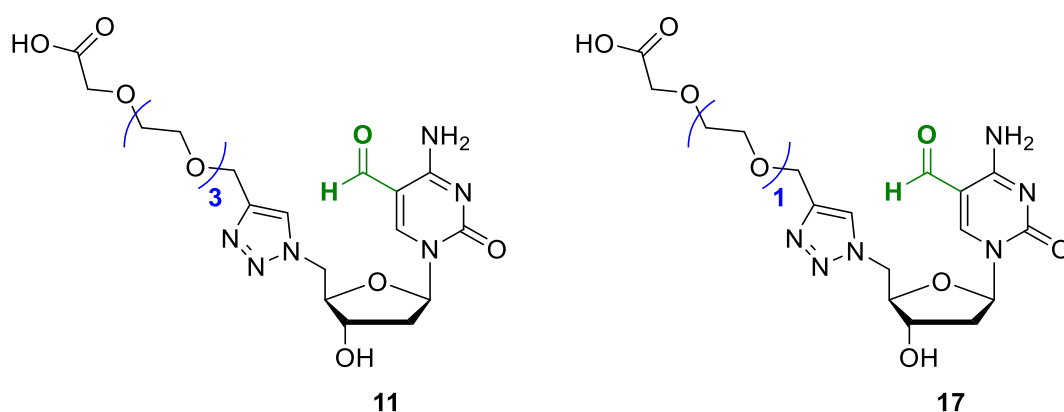


Abbildung 31: Übersicht der für die zweite Generation gewählten und verwendeten Haptene, fdC-PEG3-Hapten **11** und das kürzere fdC-PEG1-Hapten **17**

3.3.1. Synthese des PEG1-fdC Haptens

Eine Möglichkeit das fdC Hapten zu verändern, besteht wie schon zuvor erwähnt, in der Veränderung der *Linker* Länge. Hierbei entschieden wir uns für eine kürzere Variante um das Antigen besser in die Struktur des Proteins einzupassen und einen möglichen Abbau zu vermindern. Der *Linker* für die zweite Generation von Antikörpern wurde daher mit Mono-Ethylenglykol (PEG1) konzipiert. Die Synthese erfolgte analog zum PEG3 Linker in drei Schritten (Abbildung 32) ^[195]

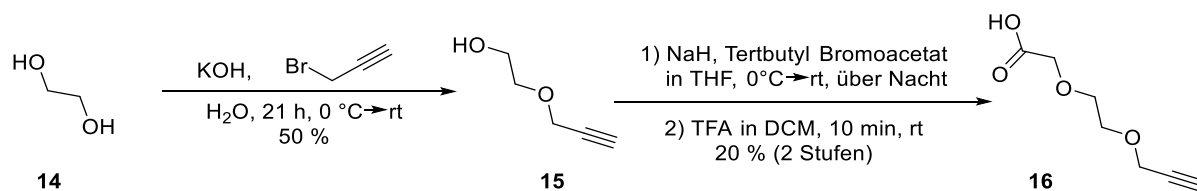


Abbildung 32: Syntheseroute zur Darstellung des PEG1-Linkers **16**, ausgehend von Ethylenglykol **14** in 2 Schritten.

Zunächst wurde Ethylenglykol **14** mit Kaliumhydroxid teilweise deprotoniert und mit Propargylbromid zum monoalkylierten Ethylenglykol **15** umgesetzt (50 % Ausbeute). Um eine Monosubstituierung zu erreichen wurde die Menge der eingesetzten Base (0.4 Äq) sowie die Menge an Propargylbromid (0.23 Äq) hinsichtlich des Ethylenglykols reduziert und langsam bei 0 °C zugetropft. Es wurde trotz dieser Bemühungen doppelte Substituierung beobachtet. Der noch freie Alkohol des monoalkylierten Ethylenglykols **15** wurde mit Natriumhydrid und *tert*-Butyl-Bromacetat umgesetzt. Nach anschließender saurer Entschützung wurde der PEG1-Linker **16** erhalten (20 % Ausbeute, über zwei Schritte) ^[195]

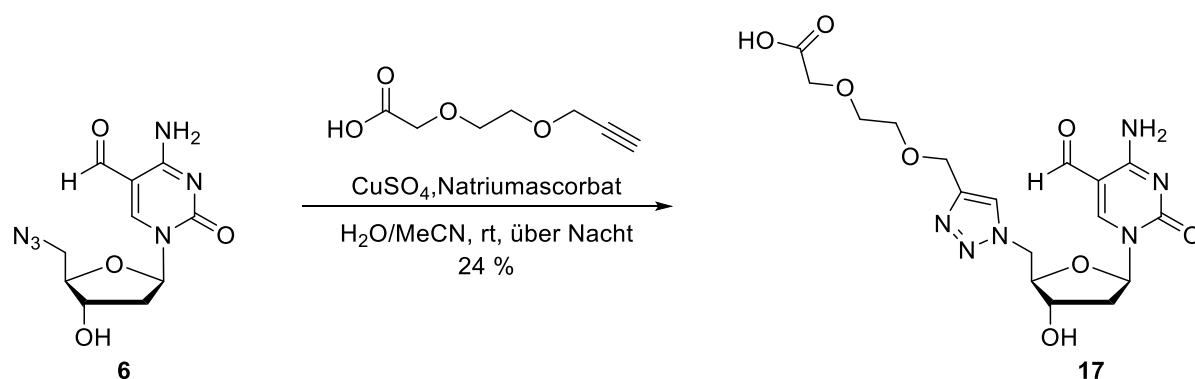


Abbildung 33: Darstellung von fdC-PEG1-Hapten **17** ausgehend von den 5'-Azid-Nukleosid **6** mittels 1,3-dipolarer Cycloaddition mit dem PEG1-Linker **16**.

Das 5'-Azid **6** und der PEG1-Linker **16** wurden mittels CuAAC verbunden (Abbildung 33). Das fdC-PEG1-Hapten **17** konnte nach zweifacher Aufreinigung mittels RP-HPLC in 24 % Ausbeute erhalten werden.

3.3.2. EDC Kopplung der Haptene **11** und **17** mit OVA und cBSA

Bei der EDC Kopplung der fdC-Haptene **11** und **17** wurde neben OVA nun auch das kommerziell erhältliche kationisierte bovine Serumalbumin (cBSA) als ein weiteres immunogenes Trägerprotein verwendet. In seiner nativen Form besitzt BSA eine Masse von 66 kDa und 59 native Lysinreste. ^[196, 197]

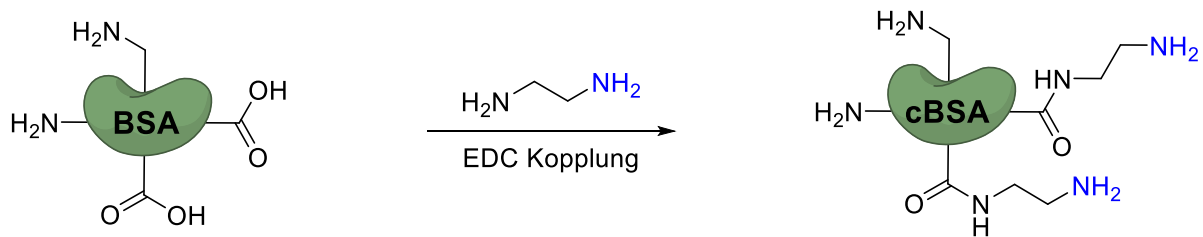


Abbildung 34: EDC-Kupplung von Ethylendiamin mit Säuregruppen des BSA, generiert durch die Peptidbindungen neue primäre Amine. Diese Amine stehen nun auch zur EDC-Kopplung mit dem in dieser Arbeit synthetisierten Hapten zur Verfügung.

Bei cBSA wurden zusätzlich die Carboxylgruppen mit einem Überschuss an Ethylendiamin und einem Carbodiimid umgesetzt. Durch die Kopplung von Ethylendiamin werden die Carboxylreste des Proteins in primäre Amine umgewandelt (Abbildung 34). Der große Vorteil den die Kationisierung des BSA mit sich bringt ist, dass nun weitaus mehr primäre Amine zur Hapten-Konjugation bereitstehen.^[198] Ein weiterer nicht zu verachtender Vorteil ist die erhöhte Immunogenität des cBSA im Vergleich zum nativen BSA.^[199] Damit ist die Wahrscheinlichkeit, dass ein Antikörper gegen das, auf dem cBSA präsentierten, Ziel-Antigen generiert wird sehr viel größer.

Das synthetisierte fdC-PEG1-Hapten **17** wurde mittels EDC Kopplung an die Trägerproteine OVA und cBSA konjugiert. Um eine möglichst große Bandbreite an Hapten-Protein-Konjugaten zur Verfügung zu haben, wurde auch das fdC-PEG3-Hapten **11** an cBSA gekoppelt (Tabelle 4). Zudem kann damit verglichen werden, ob die Länge des *Linkers* für die erfolgreiche Antikörper-Generierung ausschlaggebend ist.

Tabelle 4: Übersicht über die verwendeten fdC-Hapten-Protein Konjugate, zur Herstellung der zweiten Antikörper-Generation

Antigen	Linker	Trägerprotein
fdC	PEG1	OVA
fdC	PEG1	cBSA
fdC	PEG3	OVA
fdC	PEG3	cBSA

3.3.3. Charakterisierung der EDC Kopplungen zweiter Generation

Die erfolgreiche Bildung von Hapten-OVA-Konjugaten wurde, wie in Abschnitt 2.3.3, mithilfe von *Michael Stadlmaier* mittels Massenspektrometrie und UV/Vis Messungen qualitativ untersucht. Die Hapten-OVA-Konjugate wurden mit Trypsin verdaut und die entstandenen Peptide auf Modifizierungen untersucht. Die Daten wurden mit *Proteom Discoverer 2.2* ausgewertet und strikt mit einer *false discovery rate* (FDR) von 1 % gefiltert. Die Abbildung 35 zeigt die identifizierten Hapten-Peptid-Konjugate (grün hinterlegt), der OVA Kopplung.

	1	11	21	31	41	51	61	71	81	91
1	MGSIGAASME	FCFDVFKEK	VHHANENIFY	CPIAIMSALA	MVYLGAQDST	RTQINKVVR	DKLPFGDSI	EAQCGTSVNV	HSSLRDILNQ	ITKPNDDVYSF
101	SLASRLYAE	RYPILPEYLQ	CVKELYRGGL	EPINFQTAAD	QARELINSW	ESQTNGIIRN	VLQPSVDSQ	TAMVLVNAIV	FKGLWEKAFK	DEDTQAMPFR
201	VTEQESKPVQ	MMYQIGLFRV	ASMASEKMKI	LELPFASGTM	SMLVLLPDEV	SGLEQLESII	NFEKLTWTS	SNVMEERKIK	VYLPRMKMEE	KYNLTSVLMA
301	MGITDVFSS	ANLSGISSAE	SLKISQAVHA	AHAEINEAGR	EVVGSAGV	DAASVSEEF	ADHPFLFCIK	HIATNAVLEF	GRCVSP	

Abbildung 35: Übersicht der mit *Proteom Discoverer 2.2* identifizierten Hapten-Peptid-Konjugate der Kopplung zwischen fdC-PEG1-Hapten **17** und OVA (grün hinterlegt). Die verwendete *false discovery rate* (FDR) lag bei 1 % und die *coverage* betrug 25.4 %. M = fdC-PEG1 Modifikation an einem Lysin; C = Carbamidomethyl; O = Oxidation von Methionin. Die Modifikationen C und O wurden nur aufgeführt, wenn sie in einer mit M modifizierten Peptidsequenz auftraten.

In den Abbildungen 36 und 37 werden die identifizierten Hapten-Peptid-Konjugate (grün hinterlegt) der cBSA Kopplung gezeigt. Zusätzlich zu den Lysin Modifikationen konnten hier Modifikationen der Glutaminsäuren und Asparaginsäuren beobachtet werden.

	1	11	21	31	41	51	61	71	81	91
1	MKWVTFISLL	LLFSSAYSRG	VFRDRTHKSE	IAHRFKDLGE	EHFKGLVLIA	FSQYLQCCPF	DEHVKLVNEL	TEFAKTCVAD	ESHAGCEKSL	HTLFGDELCK
101	VASLRETYGD	MADCCKEQEP	ERNECFLSHK	DDSPDLPKLK	PDPTNLCDFF	KADEKKFWCK	YLVEIARRHP	YFYAPELLYY	ANKYNGVFQE	CCQAEDKGAC
201	LLPKIETMRE	KVLASSARQR	LRCASIQKFG	ERALKAWSVA	RLSQKFPKAE	FVEVTKLVTD	LTKVHKECCH	GDLLLECADDR	ADLAKYICDN	QDTISSKLKE
301	CCDKPLLEKS	HCIAEVERKDA	IPENLPPLTA	DFAEDKDVCK	NYQEAQDAFL	GSFLYEYSRR	HPEYAVSVLL	RLAKEYEATL	EECCAADDPH	ACYSTVFDKL
401	KHLVDEPQNL	IKQNCQDFEK	LGEYGFQNAL	IVRYTRKVPQ	VSTPTLVEVS	RLSGKVGTRC	CTKPESERMP	CTEDYLSLIL	NRLCVLHEKT	PVSEKVTKCC
501	TESLVNRRPC	FSALTPDETY	VPKAFDEKLF	TFHADICTLP	DTEKQIKKQT	ALVELLKHHP	KATEEQLKTV	MENFVAFVDK	CCAADDKEAC	FAVEGPKLVV
601	STQTALA									

Abbildung 36: Übersicht der mit *Proteom Discoverer 2.2* identifizierten Hapten-Peptid-Konjugate der Kopplung zwischen fdC-PEG1-Hapten **17** und cBSA (grün hinterlegt). Die verwendete *false discovery rate* (FDR) lag bei 1 % und die *coverage* betrug 31.8 %. M = fdC-PEG1 Modifikation an einer aminierten Glutamin-/Asparaginsäure; N = fdC-PEG1 Modifikation an einem Lysin; C = Carbamidomethyl; O = Oxidation von Methionin. Die Modifikationen C und O wurden nur aufgeführt, wenn sie in einer M/N modifizierten Peptidsequenz auftraten.

	1	11	21	31	41	51	61	71	81	91
1	MKWVTFISLL LLFSSAYSRG VFRDTHKSE IAHRFKDLGE EHFKGLVLIA FSQYLQCCPF DEHVKLVLNEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK									
101	VASLRETYGD	MADCCEKQEP	ERNECF ^{MC} LSHK	DDSPDLPLKL	PDPT ^M LCDEF	KADEKKFWGK	YLYEIARRHP	YFYAPE ^M LLYY	ANKYNGVFQE	CCQAEDKGAC
201	LLPKIETMRE	KVLASSARQR	LRCASTQKFG	ERALKAWSVA	RLSQK ^M FPKAE	FVEVTKLVTD	LTKVHKECCH	GD ^{CC} LLECADDR	ADLAKYICDN	QDTISSKLKE
301	CCDKPLLEKS	HCIAEVEKDA	IPENLPPLTA	DFAEDKDVC	NYQEAKDAFL	GSFLYEYSRR	HPEYAVSVLL	RLAKEYEATL	EECCA ^{MM} KDDPH	ACYSTVFDKL
401	KHLVDEPQNL	IKQNC ^{MM} DQFEK	LGEYGFQNAL	IVRYTRKVPQ	VSTPTLVEVS	RLSGKVGTRC	CTKPESERMP	CTEDYLSLIL	NRLCVLHEKT	PVSEKVTKCC
501	TESLVNRRPC	FSALTPDETY	VPKAFDEKLF	TFHADICTLP	DTEKQIKKQT	ALVELLKHKP	KATEEQLKTV	MENFVAFVDK	CCAAD ^{OM} KEAC	FAVEGPKLVV
601	STQTALA									

Abbildung 37: Übersicht der mit *Proteom Discoverer 2.2* identifizierten Hapten-Peptid-Konjugate der Kopplung zwischen fdC-PEG3-Hapten **11** und cBSA (grün hinterlegt). Die verwendete *false discovery rate* (FDR) lag bei 1 % und die *coverage* betrug 28,5 %. M = fdC-PEG1 Modifikation an einer aminierten Glutamin-/Asparaginsäure; N = fdC-PEG1 Modifikation an einem Lysin; C = Carbamidomethyl; O = Oxidation von Methionin. Die Modifikationen C und O wurden nur aufgeführt, wenn sie in einer M/N modifizierten Peptidsequenz auftraten.

Um die Effizienz der Konjugation zu zeigen wurde, zusätzlich zur Massenspektrometrie und den UV/VIS-Messungen, eine Natriumdodecylsulfat-Polyacrylamid-Gelelektrophorese (SDS-PAGE) durchgeführt. Vergleichbar mit einem *Electrophoretic Mobility Shift Assay* (EMSA) wird bei erfolgreicher Kopplung eine Veränderung im Laufverhalten der Proteine erwartet.

In Abbildung 38 ist die SDS-PAGE der Konjugate aus OVA und fdC-PEG1-Hapten sowie cBSA und fdC-PEG1/3-Hapten, jeweils im Vergleich zu OVA bzw. cBSA, abgebildet.

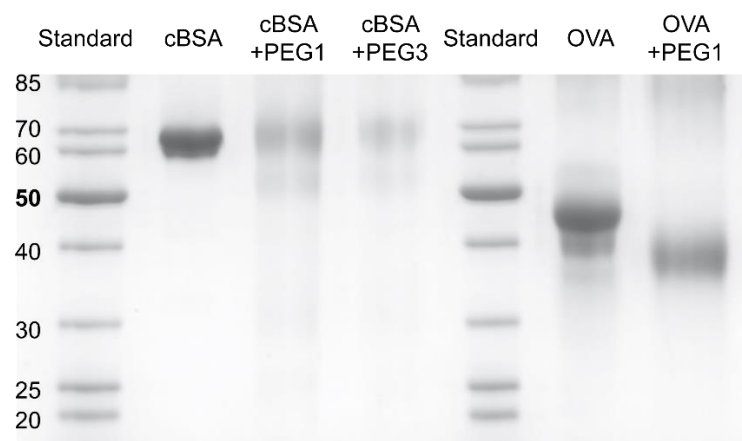


Abbildung 38: Denaturierende SDS-PAGE der fdC Konjugate von OVA und cBSA. (67 kDa)

Vergleicht man die Bande von OVA mit der des OVA-fdC Konjugats, ist eine deutliche Verschiebung um ca. 10 kDa sichtbar. Der Vergleich von cBSA mit dem konjugierten cBSA zeigt eine Verschiebung um ca. 5-10 kDa. Diese Verschiebungen deuten auf eine Abdeckung der Amine mit dem Hapten hin. Neben der veränderten Masse konnte auch ein verändertes Laufverhalten

beobachtet werden. Im Vergleich zu OVA und cBSA, zeigten die Konjugate keine distinkten Banden mehr (schmierig). Durch die durchgeführte EDC-Kopplung werden die Proteine unterschiedlich stark konjugiert und können daher keine klaren Banden mehr zeigen.

Zusammengenommen zeigten die Ergebnisse der Charakterisierung, dass von einer erfolgreichen Kopplung der fdC Haptene an die Proteine OVA und cBSA ausgegangen werden konnte. Die Konjugate wurden anschließend zur Immunisierung an das *Helmholtz Zentrum München* gegeben.

3.3.4. Optimierung der Oligonukleotide für das ELISA-Screening

Das bisherige ELISA-Screening, mit den hier synthetisierten fdC-Oligonukleotiden ODN 1-4, konnte einige potenzielle Antikörper in den Hybridomzellen-Überständen identifizieren. Jedoch waren die Menge der identifizierten Antikörper Kandidaten geringer, als beim sonst üblichen ELISA-Screening mit Hapten-Protein Bruchstücken. Zum einen könnte dies auf die stärkere Vorauswahl durch das neue Verfahrens zurückzuführen sein, als positiver Effekt. Jedoch könnten die bisher verwendeten Oligonukleotide, wider Erwarten, zu wenig Oberfläche zur Bindung der Antikörper bieten. Um diesen negativen Effekt auszuschließen, sollten längere Oligonukleotide synthetisiert werden, welche nun vier fdC-Modifikationen besitzen, um die Dichte der Modifikationen beizubehalten.

Um die Sequenzen deutlich verlängern zu können wurden keine randomisierten Sequenzen mehr verwendet. Die gewählten Oligonukleotid-Sequenzen waren ODN 10 (fdC) und ODN 7 (dC). Diese wurden so gewählt, dass sie keine Sekundärstruktur ausbilden. Außerdem sollten sie sich von den Oligonukleotid-Sequenz für den Dot Blot unterscheiden. Damit sollte weiterhin ausgeschlossen werden, dass ein potentieller Antikörper sequenzabhängig ist (Tabelle 5).

Tabelle 5: Übersicht der Sequenzen der modifizierten DNA Oligonukleotide für das ELISA-Screening zweiter Generation. Das Oligonukleotid ODN 8 wurde von *Edris Parsa* zur Verfügung gestellt.

ODN #	Sequenz
ODN 7	5'–B TTC ATC ATC GTC GTC GTC GTC ATC ATC TT–3'
ODN 8	5'–B TTC ATC ATmdC GTmdC GTmdC GTmdC GTC ATC ATC TT –3'
ODN 9	5'–B TTC ATC AThmdC GThmdC GThmdC GThmdC GTC ATC ATC TT–3'
ODN 10	5'–B TTC ATC ATfdC GTfdC GTfdC GTfdC GTC ATC ATC TT–3'
ODN 11	5'–B TTC ATC ATcadC GtcadC GtcadC GtcadC GTC ATC ATC TT–3'

Ein weiteres Problem der bisherigen Antikörper stellte die schlechte Selektivität gegenüber fdC im Vergleich zu den weiteren epigenetischen Basen dar. Um die Kandidaten für einen spezifischen fdC-Antikörper noch präziser einzugrenzen, wurden daher zusätzlich Oligonukleotide synthetisiert in denen das fdC durch mdC, hmdC bzw. cadC (ODN 8, ODN 9 und ODN 11) ersetzt wurde. Die potenziellen Antikörper können somit schon im ELISA-Screening auf mögliche Kreuzreaktivitäten zu anderen Modifikationen untersucht werden.

3.3.5. Auswertung der Dot Blots - Überstände zweiter Generation

Die Dot Blots wurden wie schon in Abschnitt 2.3.6 beschrieben durchgeführt.

Tabelle 6: Übersicht der mittels Dot Blot getesteten Hybridomzellen-Überstände der zweiten Antikörper Generation. Die Selektivität in den Dot Blots wurde durch eine Verdünnungsreihe in schwach, mittel und stark eingeordnet. DNA-E steht für die Modifikation fdC

Bezeichnung	IgG Subklasse (Ratte)	Selektivitäten (im Dot Blot)
DNA-E1 R2A 5fC 21G2	IgG2a	Keine Signale
DNA-E1 R2A 5fC 10F2	IgG2a	Keine Signale
DNA-E R2C 5fC 7H2	IgG2c	Keine Signale
DNA-E R2C 5fC 8F7	IgG2c	Keine Signale
DNA-E3 RG1 5fC 12B2	IgG1	fdC (stark), hmdC (mittel)
DNA-E3 RG1 5fC 9E5	IgG1	fdC (stark), hmdC (mittel)

Auch hier zeigten die ersten vier Chargen der zweiten Generation **21G2**, **10F2**, **7H2** und **8F7** keine Bindung an die auf der Nylon Membran immobilisierten Oligonukleotide. (siehe Tabelle 6)

Erfreulicherweise zeigten die Chargen **12B2** und **9E5** hingegen jeweils eine gute Selektivität gegen fdC (Abbildung 39).

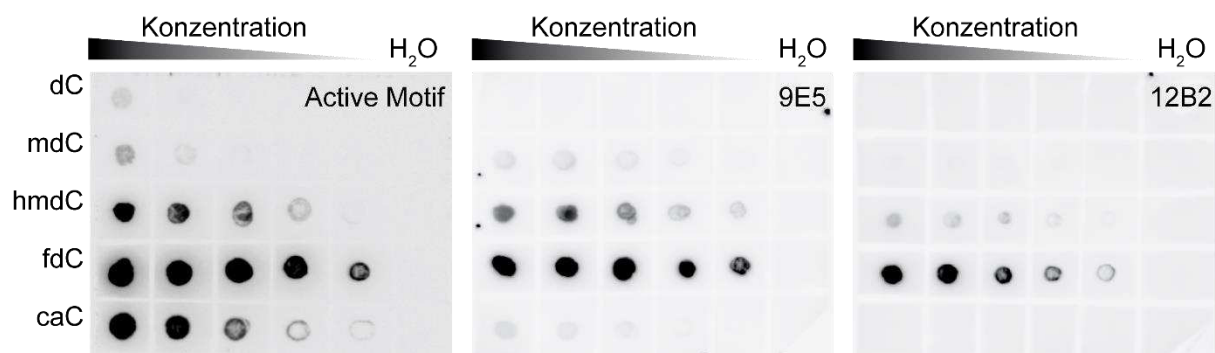


Abbildung 39: Dot Blots des polyklonalen Antikörpers (*Active Motif*) gegen fdC und die Hybridomaüberstände 9E5 und 12B2. Die Oligonukleotide (Sequenz C) mit den unterschiedlichen Modifikationen (dC, mdC, hmdC, fdC und caC) wurden ausgehend von 21 μ M (erste Spalte) jeweils um die Hälfte verdünnt aufgetragen. In der letzten Spalte wurde ddH₂O als Negativkontrolle aufgetragen.

Beide Antikörper **12B2** und **9E5** besitzen dabei keine Bindung gegenüber dem dC Oligonukleotid. Wohingegen der kommerziell erhältliche polyklonale Antikörper (*Active Motif*) bei erhöhter Konzentration eine Affinität zu dC aufweist. Die Modifikation mdC werden vom *Active Motif* Antikörper und dem **9E5** Überstand nur bei sehr hohen Konzentrationen des Oligonukleotids gebunden. Erstaunlicherweise zeigt **12B2** keinerlei Bindung gegenüber der mdC Modifikation. Vor allem hinsichtlich der Abundanz von mdC in genomischer DNA, kann dies ein großer Vorteil sein. Alle drei

Antikörper besitzen eine Bindung gegenüber hmdC, wobei diese bei **12B2** am schwächsten ausfällt. Der größte Unterschied zwischen dem polyklonalen Antikörper von *Active Motif* und den Überständen **12B2** und **9E5** zeigt sich bei der Modifikation cadC. Hier bindet der *Active Motif* Antikörper in zu hmdC vergleichbarer Stärke. **12B2** und **9E5** zeigen jedoch kaum bzw. keine Bindung.

Die sehr starke Bindung von **12B2** und **9E5** an fdC überstiegen das im Dot Blot gewählte Konzentrationsgefälle. Daher wurde eine weitere Konzentrationsreihe mit einer geringeren Endkonzentration von 0.15 μM gewählt, um ein Minimum der Bindung zu erreichen.

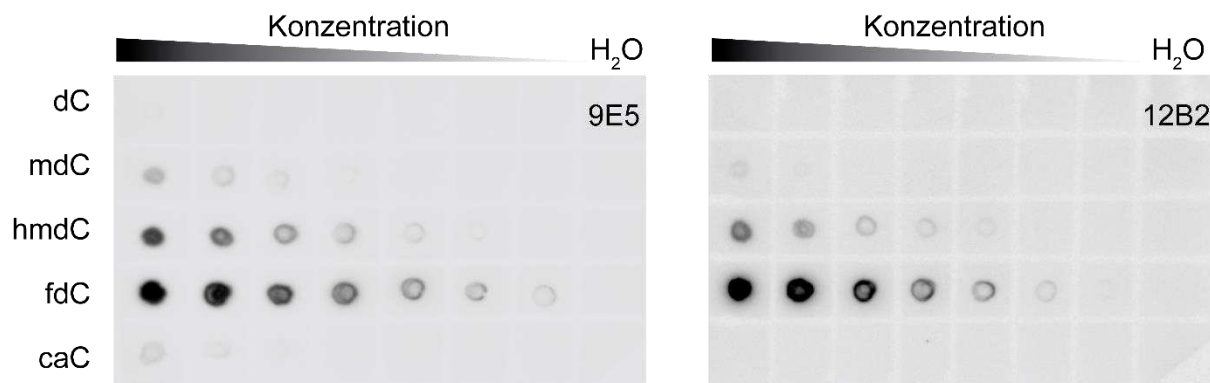


Abbildung 40: Dot Blot der Hybridoma-Überstände 12B2 und 9E5 mit erweiterter Konzentration. Die Oligonukleotide mit den unterschiedlichen Modifikationen (dC, mdC, hmdC, fdC und caC) wurden ausgehend von 21 μM (erste Spalte) jeweils um die Hälfte verdünnt aufgetragen (bis 0.15 μM). In der letzten Spalte wurde ddH₂O als Negativkontrolle aufgetragen.

In den Dot Blots mit erweiterter Konzentrationsreihe konnten die bisherigen Ergebnisse bestätigt werden (Abbildung 40). Besonders bei den Oligonukleotid Konzentrationen von 0.3 μM und 0.15 μM kann bei beiden Überständen (**12B2** und **9E5**) keine Bindung an hmdC beobachtet werden, bei gleichzeitiger Bindung an fdC. Das zeigt, dass beide Überstände das fdC Antigen bevorzugen.

Das *Helmholtz Zentrum München* führte zur weiteren Charakterisierung einen *Capture ELISA* mit **12B2** und **9E5** durch und konnte damit die Fähigkeit der Antikörper zur Immunopräzipitation (IP) testen. Im *Capture ELISA* wird eine IP simuliert, indem die Antikörper immobilisiert werden und nicht die Oligonukleotide. Die Antikörper werden daraufhin mit Antigen enthaltenden Oligonukleotiden inkubiert und die Menge, welche gebunden werden konnte, untersucht. Dabei konnte mit dem Hybridoma-Überstand **9E5** eine gute und fdC spezifische IP gemacht werden. Mit dem Hybridoma-Überstand **12B2** konnte jedoch keine IP erreicht werden, ist aber weiterhin für nicht-IP Anwendungen, wie die Immunfluoreszenz, interessant. Daher wurden nur die Antikörper aus dem Hybridoma-Überstand **9E5** mittels *fast protein liquid chromatography* (FPLC) gereinigt. Die Antikörper Konzentration im Überstand lag bei 100 $\mu\text{g/mL}$. Der Hybridoma-Überstand **9E5** und der gereinigte Antikörper befindet sich momentan für weitere Tests zur Sequenzierungsmethoden-Entwicklung mit genomischer DNA bei der *Arbeitsgruppe Epigenetik* der *Universität des Saarlandes* (Prof. Dr. Jörn

Walter). Dabei soll, basierend auf dem hier generierten Antikörper, eine Anreicherung der fdC Modifikation in genomischer DNA erreicht werden. Damit könnte eine anschließende Sequenzierung sehr viel präzisere Ergebnisse liefern. Abschließende Ergebnisse stehen jedoch noch aus.

3.4. Anwendungsmöglichkeiten der neuartigen Methode

Mit der von mir entwickelten neuartigen Methode konnten erstmals monoklonale Antikörper gegen fdC hergestellt werden. Diese sind mit der Affinität des kommerziell erhältlichen polyklonalen fdC-Antikörpers von *Active Motif* vergleichbar. Zudem ist die Kreuzreaktivität bezüglich dC und cadC im Vergleich zum kommerziellen Produkt deutlich geringer. Die Etablierung dieser Methode ermöglicht es nun, die Herstellung von Antikörpern auf weitere Modifikationen auszuweiten, welche durch die Periodatspaltung nicht zugänglich sind.

Ein geeignetes Anwendungsbeispiel wäre die RNA Hypermodifikation Queuosin (Q) (Abbildung 41). Diese Hypermodifikation hat in letzter Zeit stark an Bedeutung gewonnen.^[200, 201, 202] Es konnte gezeigt werden, dass die C38-Methylierung in tRNA^{Asp} durch Dnmt2 ohne die Anwesenheit von Q nicht stattfinden kann.^[203] Queuosin wird zudem mit dem Schutz vor Ribonuklease-Abbau und mit der Proteintranslation in Verbindung gebracht.^[200, 202] Bisher existieren jedoch keine Antikörper gegen Q. Da Q einen Cyclopentendiol-Rest als Basen-Modifikation aufweist, ist eine Periodat-basierte Antikörper-Entwicklung nicht möglich: Das Periodat würde nicht nur die Ribose, sondern auch die Modifikation selbst spalten.^[204] Auch für die verwandten Modifikationen manQ und galQ wäre eine Periodat-basierte Antikörper-Entwicklung ausgeschlossen. Die von mir entwickelte Methode bietet nun erstmals die Möglichkeit, Antikörper gegen Q, manQ und galQ zu generieren.

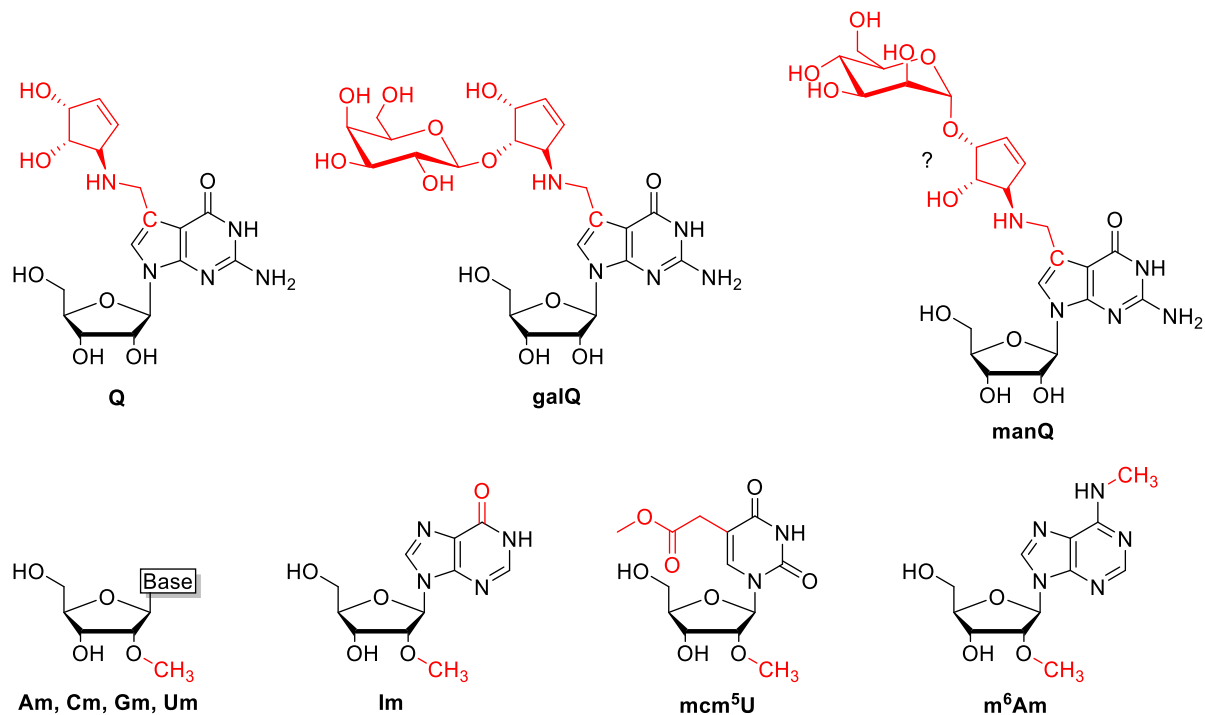


Abbildung 41: Auswahl an anspruchsvollen Modifikationen, gegen die, mit der in dieser Arbeit etablierten Methode, nun erstmals/spezifischere Antikörper hergestellt werden könnten. Das Nukleosid-Grundgerüst ist schwarz dargestellt und die Modifikation in rot. Alle hier dargestellten Modifikationen kommen in RNA von Eukaryoten vor. Die tatsächliche Struktur von manQ ist noch nicht abschließend geklärt (schwarzes Fragezeichen).

Neben dem Bereich der Hypermifikationen sind auch Antikörper gegen einfachere Modifikationen, wie Methylierungen, von großem Interesse. Aktuell werden stetig neue Methylierungen in mRNA gefunden: *N*⁶-Methyladenosin (m⁶A),^[125, 126] *N*¹-Methyladenosin (m¹A),^[205, 206] 2'-*O*-Methylierungen der Ribose (Nm)^[207] etc. Die meisten dieser Modifikationen werden über Antikörper angereichert und sequenziert. In vielen Fällen wird kontrovers diskutiert, ob die Kreuzreaktivitäten der verwendeten Antikörper eine Anreicherung unspezifischer Sequenz-Motive begünstigen.^[153, 168] Beispielsweise gibt es Hinweise darauf, dass die Sequenzierungsergebnisse von m¹A, auf eine Kreuzreaktivität des m¹A-Antikörpers mit der *N*⁷-Methylguanosin (m⁷G) Modifikation zurückzuführen sind.^[208] Ähnliches gilt für m⁶A, dessen Sequenzierungsergebnisse teilweise mit einer Kreuzreaktivität des verwendeten Antikörpers mit *N*⁶-2'-*O*-Dimethyladenosin (m⁶Am)^[154] in Verbindung gebracht werden.

Der große Vorteil der hier vorgestellten Methode ist, dass das gesamte Nukleosid intakt bleibt und nicht nur die Base. Daher sollten Antikörper hergestellt werden können, welche sowohl eine Modifikationen der Basen als auch der Ribose spezifisch erkennen. Ein Antikörper der zwischen m⁶A und m⁶Am unterscheiden kann wird damit möglich.^[168] Dies würde eine entscheidende und unbedingt notwendige Verbesserung der Sequenzierung mittels Antikörper-Anreicherung ermöglichen.

4. *Hairpin* Oligonukleotide für die Sequenzierung von mdC, hmdC, fdC und cadC in dsDNA

4.1. Design der *Hairpin* Oligonukleotide

Die *Hairpin* Oligonukleotide wurden nach dem in Abbildung 42 gezeigten Grundgerüst konzipiert, welches sich in drei Hauptbereiche aufteilt: Dem *loop*-Bereich mit der Barcode Sequenz, dem Stammbereich mit den vier kanonischen Basen und den epigenetischen Basen, sowie dem nukleasespezifischen Überhang am 5'-Ende.

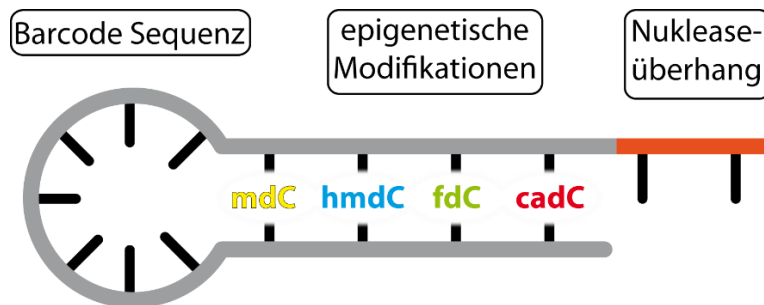


Abbildung 42: Schematische Darstellung der drei Hauptbereiche der *Hairpin* Oligonukleotide. Der *loop*-Bereich beinhaltet eine randomisierte Sequenz, welche als Barcode fungiert. Alle vier epigenetischen Basen sind in der Stammregion untergebracht. Einen Nuklease spezifischen Überhang bildet das Ende des *Hairpin* Oligonukleotids.

Der Stammbereich der *Hairpins* ist jeweils gleich aufgebaut. Er besitzt eine klar definierte Sequenz. Die Sequenz beinhaltet, neben den vier epigenetischen Basen, drei unmodifizierte Cytosine, sowie zwei A-T Paare (siehe Tabelle 7). Die definierten Cytosine (modifiziert und unmodifiziert) sollen später dazu dienen, die Qualität der erfolgten Bisulfit-Behandlung zu beurteilen (Abbildung 43).

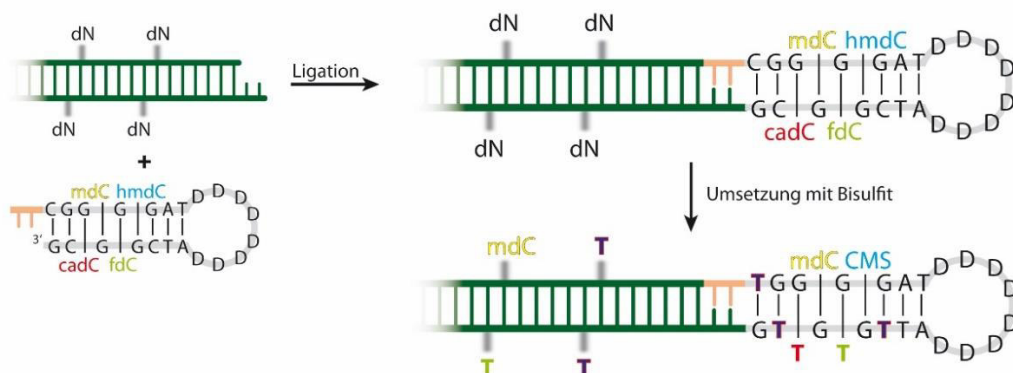


Abbildung 43: Das *Hairpin* Oligonukleotid als interner Standard zur Überprüfung der Bisulfit-Reaktion. Nach der Ligation mit der nukleasespezifisch geschnittenen genomischen dsDNA und der Umsetzung mit Bisulfit, kann mittels den bekannten Modifikationen eine durchschnittliche Bisulfit-Umsetzungsrate ermittelt werden. In dieser Abbildung ist ein Umsatz von 100% gezeigt.

Nach einer erfolgreichen und vollständigen Reaktion mit Bisulfit sollten die unmodifizierten Cytosine als Thymin sequenziert werden. Eine Abweichung zeigt wie solide die erhaltenen Daten sind. Die vier modifizierten Cytosine sollten je nach Sequenzierungsmethode als Thymin bzw. Cytosin gelesen werden und liefern damit eine Aussage über die Effektivität der Umwandlungsmethode und gleichzeitig über die Reaktion mit Bisulfit.

Die Sequenzen der in dieser Arbeit, in Zusammenarbeit mit *Martin Rossa* synthetisierten *Hairpins* sind in Tabelle 7 zusammengefasst.

Tabelle 7: Sequenzen aller in dieser Arbeit synthetisierten *Hairpin* Oligodesoxyribonukleotide. Die entsprechenden Restriktionsenzyme und deren jeweiligen Schnittmotiv, welche den gewünschten Überhang erzeugen; W = A/T; D = A/G/T; N = A/C/G/T

ODN #	Sequenz in hybridisierter Form	Restriktionsenzym; Schnittmuster
ODN 12		MspI 5' - C C G G - 3' 3' - G G C C - 5'
ODN 13		DdeI 5' - C T N A G - 3' 3' - G A N T C - 5'
ODN 14		BsaWI 5' - C C G G - 3' 3' - G G C C - 5'
ODN 15		Eco47I 5' - G G W C C - 3' 3' - C C W G G - 5'
ODN 16		A-Tailing

Der *loop*-Bereich des *Hairpins* (Barcode Sequenz) soll mit den kanonischen Nukleobasen dA, dG und T teilrandomisiert synthetisiert werden (D= dA, dG oder T). Die Nukleobase dC soll dabei bewusst weggelassen werden, um keine starken C-G Bindung zu ermöglichen, welche die Ausbildung des gewünschten Loopbereichs, des Stammbereichs und des Nuklease-Überhangs behindern könnte. Die randomisierte Barcode Sequenz hat einen entscheidenden Vorteil gegenüber einem definierten *loop*-Bereich. Die durch den *Hairpin* kovalent verbundene DNA muss vor der Sequenzierung durch PCR amplifiziert werden. Um das bei der Amplifizierung auftretende Risiko von Produkt Redundanz und

Kontaminierung zu minimieren, kann der randomisierte *loop*-Bereich des *Hairpin* als einzigartiger Barcode fungieren. Der Barcode kann dafür verwendet werden die gewünschten Sequenzen zu markieren, sowie redundante und kontaminierende Sequenzen auszuschließen.^[209]

Um die dsDNA mit den *Hairpin* Oligonukleotide ligieren zu können, muss ein DNA Doppelstrang-Bruch mit einem spezifischen Überhang generiert werden. Dies geschieht mit Nukleasen, die wiederum ein distinktes Sequenzmuster erkennen und schneiden. Allgemein sollten die Schnittpunkte der Nukleasen und damit die Positionierung der *Hairpin* Oligonukleotide so gewählt werden, dass sie möglichst nahe an der zu untersuchenden Stelle des Genoms liegt. *Giehr et al.* schlagen vor, dass die Entfernung zur Region von Interesse nicht mehr als 250 Basenpaare beträgt.^[172] Um daher möglichst viele Nukleaseschnittpunkte im Genom abzudecken, sollten fünf *Hairpin* Oligonukleotide mit, sich jeweils unterscheidende, Überhängen am 5'-Ende ausgestattet werden. Die fünf konzipierten Überhänge passen zu den Schnittmotiven, unterschiedlicher „Typ II“-Restriktionsenzymen. Die Restriktionsenzyme wurden so gewählt, dass deren Aktivität möglichst nicht durch vorhandene epigenetische Modifikationen beeinträchtigt wird. Andernfalls könnten stark modifizierte Regionen nicht geschnitten werden.

Der *Hairpin* ODN **12** besitzt beispielsweise einen 5'-CG Überhang, welcher zum Schnittmotiv, des Restriktionsenzym *MspI*, passt. *MspI* ist ein bekannter Vertreter der Restriktionsenzyme und findet in mdC-Sequenzierungsmethoden Verwendung.^[100, 101] *MspI* macht dabei keinen Unterschied, ob die Erkennungsstelle 5'-CCGG-3' am zweiten C methyliert ist.^[210] Der teilrandomisierten 5'-TNA Überhang von ODN **13**, ist das Gegenstück zum Schnittmotiv 5'-CTNAG-3' der Endonuklease *DdeI* (N = A,C,G und T).^[211] ODN **14** weist hingegen einen 5'-CCGG Überhang auf, um dem Schnittmotiv 5'-WCCGGW-3' (W = A, T) von *BsaWI* zu entsprechen.^[212] Sowohl *DdeI*, als auch *BsaWI* schneiden ihre Schnittmotive unabhängig vom epigenetischen Modifizierungsgrad.^[213] Das ODN **15** besitzt wie ODN **13** einen teilrandomisierten Überhang 5'-GWC, um dem Schnittmotiv 5'-GGWCC-3' von *Eco47I* zu entsprechen (W = A, T). *Eco47I* besitzt aber den Nachteil, dass es die Sequenz bei einer Methylierung des zweiten Cytosin nicht schneiden kann.^[213, 214] ODN **16** hat, im Gegensatz zu den anderen *Hairpins*, keinen 5' sondern einen 3'-T Überhang. Dieser Überhang ist für dsDNA mit einem *blunt end*, also für einem Doppelstrang ohne Überhang, konzipiert. Mithilfe verschiedener Methoden,^[215] wie dem Klenow Fragment (DNA Polymerase ohne 5' nach 3' Exonuklease Aktivität),^[216, 217] oder der Polymerase Typ- α aus Hühnerembryos,^[218] kann ein einzelner A-Überhang an der *blunt end* dsDNA erzeugt werden. Dieser lässt sich dann mit dem T-Überhang des *Hairpin* ligieren.^[172] Der Vorteil hier ist, dass man *blunt ends* unabhängig vom eingesetzten Restriktionsenzym, z.B. mittels der ssDNA selektiven *mung bean nuclease*,^[219] erzeugen kann. Eine weitere Möglichkeit bietet der dsDNA-Bruch mittels Ultraschall, dieser zeigt aber Sequenzabhängigkeit.^[220]

4.2. Synthese und Charakterisierung der *Hairpin* Oligonukleotide

Durch *Schröder et al.* konnte gezeigt werden, dass die Synthese eines Oligonukleotid mit allen vier Cytosin Modifikationen möglich ist.^[221] Trotz der, im Gegensatz zu den kanonischen Basen, geringeren Einbaueffizienz der modifizierten Basen. Jedoch stellt die *Hairpin*-Struktur, aufgrund der Ausbildung der Sekundärstruktur, auch während der Festphasensynthese, eine zusätzliche Herausforderung dar. Diese sehr viel komplexere Struktur konnte nur mit weiterer Optimierung der Synthese möglich gemacht werden. Um die Effizienz zu erhöhen, wurden die modifizierten Phosphoramidite mit einer höheren Konzentration von 125 mM-150 mM gelöst und die Kopplungszeit auf 3 min erhöht (Konzentration kanonischer Basen = 100 mM). Neben der verlängerten Kopplungszeit wurden nach 150 s die Reaktionsmischung 0.1 s über die Kartusche hin und zurück bewegt (8 x mit jeweils 4 s Reaktionszeit), um eine Durchmischung zu gewährleisten. Die kanonischen Phosphoramidite wurden nach dem Standardprotokoll mit einer Konzentration von 100 mM gelöst und mit den sonst üblichen 45 s gekoppelt. Die Oligonukleotide besitzen zudem unterschiedliche Randomisierungen, was besonders die Charakterisierung erschwert. Für den Einbau von A/T (W), A/G/T (D) und A/C/G/T (N) wurden, wie in Abschnitt 2.3.5, jeweils die entsprechenden kanonischen Nukleosidphosphoramidite, bei einer gesamt Konzentration von 100 mM, äquimolar vorgemischt. Diese vorgemischte Lösung wurde dann im Standardprotokoll in der Festphasensynthese eingesetzt.

Eine weitere Herausforderung stellte die Ausbildung der *Hairpin*-Strukturen schon während der Synthese dar. Als Trägermaterial wurde daher *controlled pore glass* (CPG), mit einer Porengröße von 1000 Å gewählt. CPG bietet eine höhere Hydrophilie und eine höhere Festigkeit, als üblicherweise verwendetes Polystyrol und wird erfolgreich zur DNA-Festphasensynthese komplizierte sekundär-Strukturen, wie beispielsweise G-Quadruplexe, verwendet.^[222] Des Weiteren wurde besonders auf eine gute Performance des DNA-Synthesizers Wert gelegt. Um diese zu überprüfen, wurden kurz vor der eigentlichen Synthese größere Test-Oligonukleotide synthetisiert. Dabei wurden die kanonischen Phosphoramidite jeweils dreimal hintereinander eingebaut, um die Stabilität der Trityl-Werte zu überprüfen. Nur wenn diese, bei Rohwerten von 500, um weniger als 5 Punkte schwanken wurde die Synthese durchgeführt (bei geringeren Anfangswerten, war die Toleranz dementsprechend geringer). Aufgrund dieser erfolgreich optimierten Synthese wurden kaum Abbruchfragmente erhalten (Abbildung 44)

Das Phosphoramidit von mdC war kommerziell erhältlich. Die weiteren modifizierten Phosphoramidite wurden nach literaturbekannten Syntheserouten, von *Kristof Hufnagel* und in Teilschritte von mir, in guten Ausbeuten synthetisiert.^[221] Die synthetisierten Oligonukleotide wurden anschließend, in Zusammenarbeit mit *Martin Rossa*, vom Trägermaterial basisch abgespalten und dabei die basenlabilen Schutzgruppen entfernt. Nach einer kurzen HPLC Aufreinigung wurden die säurelabilen Schutzgruppen von den Oligonukleotiden entfernt. Aufgrund der randomisierten Sequenzen, war der

Reaktionsfortschritt nur schwer zu erkennen, konnte aber mit Hilfe von zeitversetzten MALDI-ToF Messungen erreicht werden.

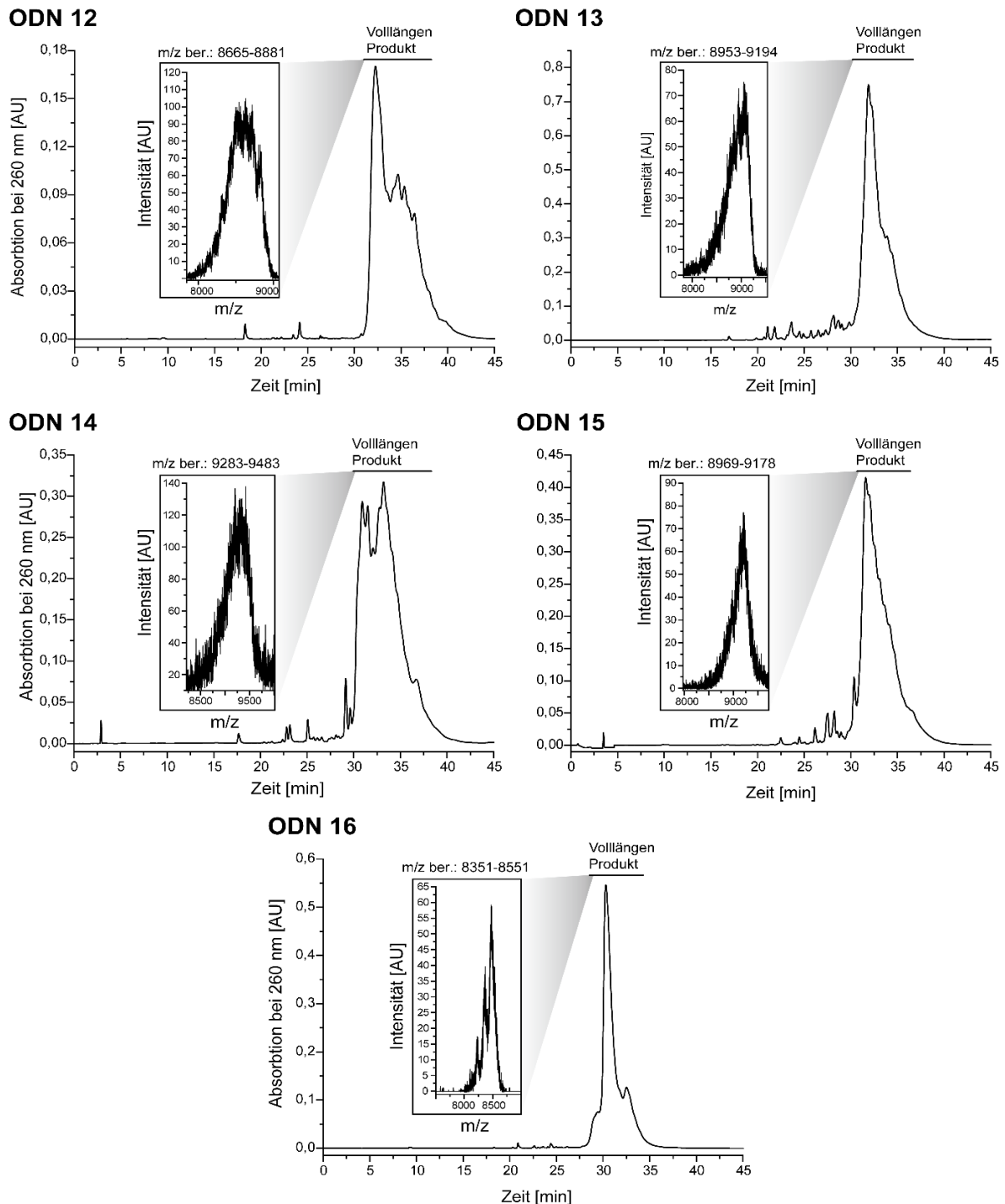


Abbildung 44: Charakterisierung der *Hairpin* Oligonukleotide ODN 12-16 mit den *crude* HPLC Chromatogrammen (großes Chromatogramm) und den dazugehörigen MALDI-ToF Spektren der gereinigten Oligonukleotide (kleines Spektrum im hervorgehobenen Kasten). Über den MALDI-ToF Spektren sind die berechneten Bereiche der möglichen Massen der Vollängenprodukte zu sehen.

In Abbildung 44 ist die Charakterisierung der *Hairpin* ODN 12-16 zu sehen. Die HPLC Chromatogramme, direkt nach der vollständigen Entschützung, zeigen aufgrund der randomisierten

Sequenzen ein sehr stark verbreitetes Produktsignal. Dieses wurde in kurzen Zeitintervallen fraktioniert, mittels MALDI-ToF untersucht und die Vollängenprodukt enthaltenden Fraktionen vereinigt.

Naturgemäß ist bei den randomisierten ODNs in der MALDI-ToF Analyse kein klares Signal zu erkennen. Es wird vielmehr eine Verteilung um den Mittelpunkt der maximal bzw. minimal berechneten Massen erwartet. Trotz der sehr anspruchsvollen Synthese konnten die *Hairpins* ODN **12-16** in guter Ausbeute und Reinheit erhalten werden. Da der Verlust durch die Analyse möglichst gering gehalten werden sollte, wurden die MALDI-ToF Spektren nahe am *Limit of Detection* (LoD) aufgenommen. In der abschließenden Charakterisierung sollte gezeigt werden, dass die Schutzgruppen vollständig abgespalten wurden. Kleinere Fragmente sollten in der weiteren Benutzung nicht stören, da sie nicht an die DNA ligiert werden können bzw. in der Sequenzierung erkannt werden können.

Bei ODN **12-16** konnte eine vollständige Entfernung der Schutzgruppen beobachtet werden. Auch hier war die Analyse der randomisierten Sequenzen sehr herausfordernd. Am Beispiel von ODN **12** wurde die leichteste Masse ($m/z = 8665$) und die schwerste mögliche Masse ($m/z = 8881$) errechnet. Das ODN **12** zeigt eine maximale Intensität um den Bereich $m/z = 8700-8800$ und liegt damit im errechneten Bereich. Auch für die ODN **13-16** konnten eine maximale Intensität in den errechneten Mittelpunkte 9074 (ODN **13**), 9383 (ODN **14**), 9074 (ODN **15**), 8451 (ODN **16**) dieser Maximal- und Minimal-Werte beobachtet werden (Abbildung 44).

Die erfolgreich synthetisierten Oligonukleotide ODN **12-16** wurden zuletzt entsalzt und lyophilisiert.

4.3. Simultane Sequenzierung der epigenetischen Basen in dsDNA

Mit den hier synthetisierten *Hairpins* sollte es nun möglich sein, alle vier Modifikationen des Cytosin gleichzeitig zu untersuchen.

Nach der Behandlung der DNA, mit einem der erwähnten Restriktionsenzyme und der Ligation des spezifisch entstehenden Überhangs, mit dem entsprechenden *Hairpin* Oligonukleotid kann die dsDNA zusammengefügt werden. Durch eine Aufteilung der Reaktion in vier Teile und einer folgenden Sequenzierung von mdC, hmdC, fdC bzw. cadC, ist eine simultane Bestimmung der Methylierungsmuster und dem Oxidationsstatus eines dsDNA Abschnitts von Interesse möglich (Abbildung 45).

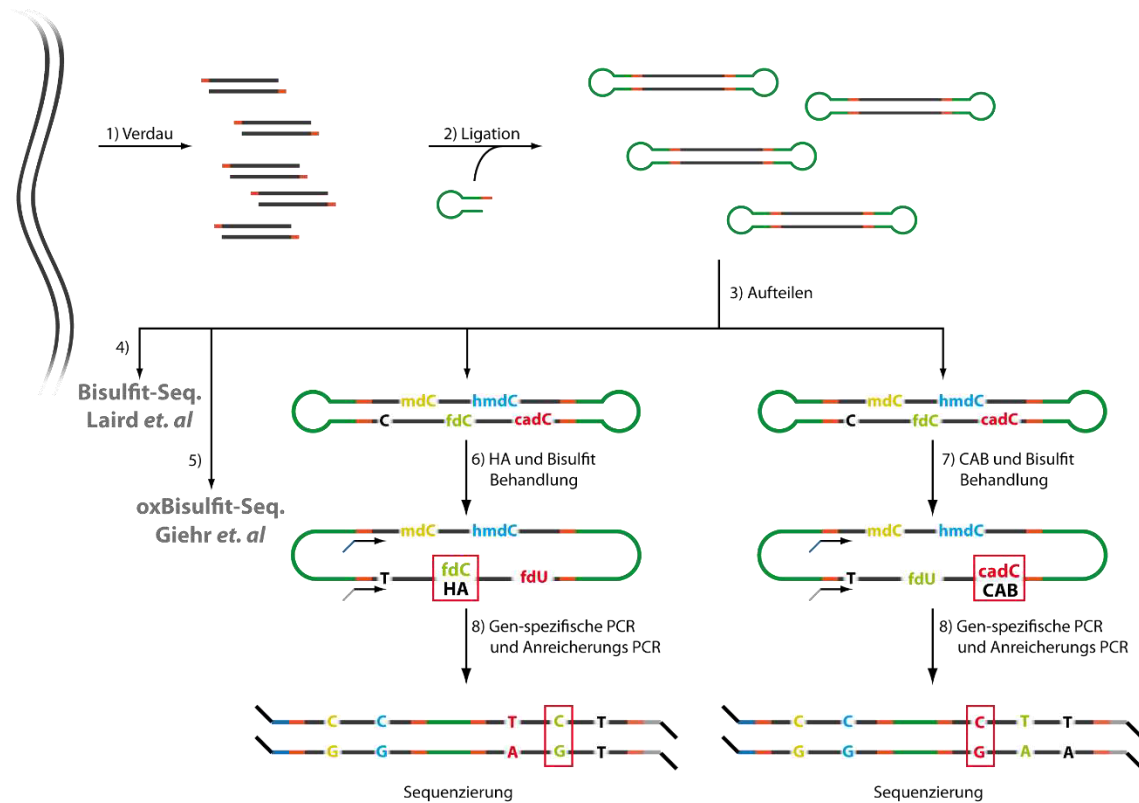


Abbildung 45: Experimenteller Aufbau der *Hairpin* Methode 1) Verdau genomischer DNA mit einer spezifischen Endonuklease. 2) Ligation der synthetisch erstellten *Hairpins* mit entsprechendem 5'-Überhang und damit kovalenter Verbindung der dsDNA. 3) Aufteilen der ligierten DNA in vier unterschiedliche Reaktionen: 4) BS Behandlung; 5) oxBS Behandlung; 6) Umsetzung von fdC mit Hydroxylamin und darauf folgende BS Behandlung; 7) *Chemical Modification-Assisted Bisulfite* (CAB) Behandlung. 8) Amplifizieren der *region of interest*, die Einführung der Sequenzieradapter und im Weiteren Schritt eine Multiplex Anreicherung mittels PCR- und ID-tagging. Die Abbildung ist an die Vorlage von Referenz ^[172] angelehnt und erweitert.

Um die *Hairpin* BS Methode erfolgreich auf die Erkennung von fdC auszuweiten, soll in Vorversuchen ein neues Hydroxylamin „SuperFly“ (SF) **18** an fdC getestet werden, welches von *Rahimoff et al.* für die Untersuchungen an abasischen Stellen entwickelt wurde (Abbildung 46).^[223]

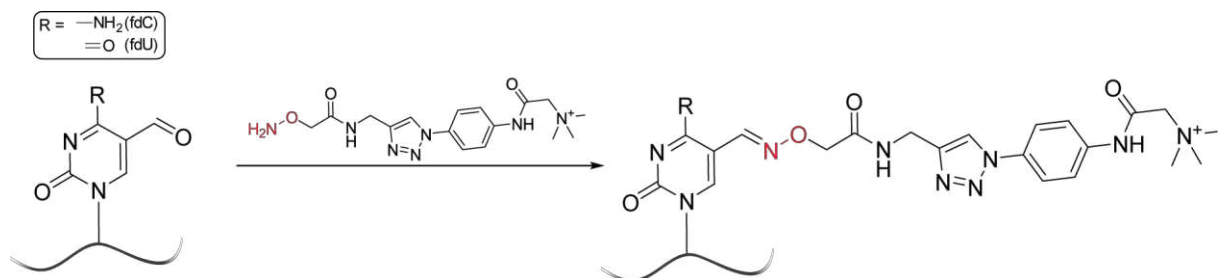


Abbildung 46: Derivatisierung der Formyl-Gruppe des fdC bzw. des fdU in DNA mit dem Hydroxylamin-Reagenz **18**.

Die durch das Hydroxylamin **18** geschützte fdC Modifikation, sollte, wie beim fCAB-seq,^[89] durch Bisulfit nicht mehr deformyliert werden können. Eine weitere Methode, welche durch die Derivatisierung mit dem Hydroxylamin möglich sein sollte, ist eine Variante des *TET-assisted BS-seq*

(TAB-seq).^[87, 224] Hier verhindert ein an hmdC gekoppelter Zucker die Umsetzung mit TET1 zu cadC. Auch das mit dem Hydroxylamin verknüpfte fdC sollte die Oxidation durch TET1 zu cadC verhindern können. Somit würden alle anderen 5-modifizierten Cytosine zu cadC oxidiert, und nur fdC im BS-seq als C ausgelesen werden.

4.3.1. Vorversuche zur Reaktivität von *SuperFly* mit fdC und fdU

Um zu zeigen, dass das sterisch anspruchsvolle Hydroxylamin **18** fdC binden kann, wurden zeitaufgelöste HPLC Studien zum Umsatz durchgeführt. Eine weitere interessante Modifikation der DNA mit freier Formyl-Gruppe ist das 5-Formyluridin (fdU). Um die beiden Modifikationen und die abasischen Stellen unterscheiden zu können, wurde fdU vergleichend zu fdC untersucht. Die Reaktivität hinsichtlich der abasischen Stelle, für die das Reagenz ursprünglich entwickelt wurde, ist ausreichend untersucht und wurde daher als Positivkontrolle verwendet. Das Reagenz reagiert innerhalb von 40 min quantitativ mit einem Oligonukleotid mit abasischer Stelle.^[223]

Für den Vergleich der Derivatisierungszeiten von fdC und fdU wurden synthetische Oligonukleotide mit der Sequenzen 5'–GTA ATG **ZGG** TAT G–3' (Z = fdC bzw. fdU) verglichen, welche von *Iacovos Michaelides* zur Verfügung gestellt wurden. Die Oligonukleotide wurden mit dem Hydroxylamin Reagenz in einer gepufferten Lösung mit bzw. ohne *p*-Anisidin als Katalysator inkubiert. Zu bestimmten Zeitpunkten wurden die Reaktionen durch die Zugabe von Aceton und durch Schockgefrieren, mittels flüssigem N₂, abgebrochen. Das überschüssige Hydroxylamin reagiert dabei mit dem in großen Mengen zugeführtem Aceton ab und wird somit aus der Reaktion schlagartig entfernt. Damit konnten auch die sehr kurzen Reaktionszeiten von 5 min und 15 min realisiert werden.

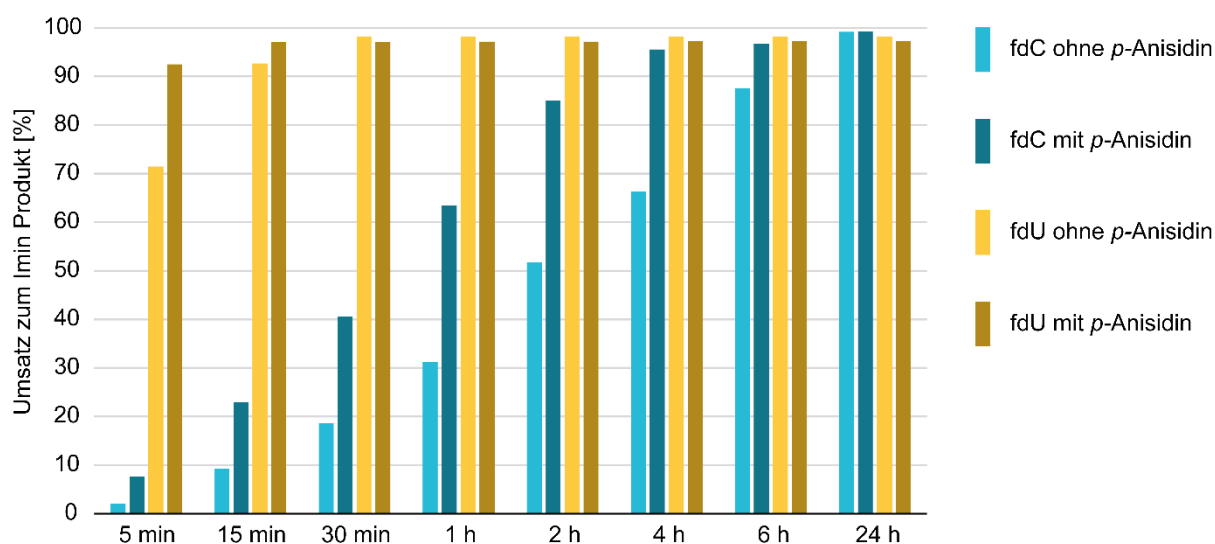


Abbildung 47: Zeitaufgelöster Vergleich des Umsatzes vom Hydroxylamin mit fdU und fdC. Die Reaktion wurde bei 37 °C durchgeführt. Die hellblauen Balken zeigen den Umsatz von fdC in Abwesenheit des Katalysators *p*-Anisidin, die dunkelblauen zeigen den Umsatz von fdC in Anwesenheit des Katalysators. Die gelben Balken zeigen den Umsatz von fdU in Abwesenheit des Katalysators, die dunkelblauen zeigen den Umsatz von fdU in Anwesenheit des Katalysators.

Der Umsatz der Reaktion konnte mithilfe von analytischer RP-HPLC nachvollzogen werden, da das gebundene Hydroxylamin Reagenz **18** eine Verschiebung in der Elutionszeit von ca. 2 min verursacht. Vergleichbar zu der später näher beschriebenen Methode zur Oxidation von hmdC (Abschnitt 5.1.2), wurden die Integrale der Signale verglichen und daraus der Umsatz abgeschätzt (Abbildung 47).

Wie erwartet reagierte auch das sterisch anspruchsvolle Hydroxylamin **18**, mit den Aldehyden von fdC und fdU. Im Einklang mit der Literatur ^[225] reagierte das Reagenz dabei sehr viel schneller mit fdU (Vollständiger Umsatz nach 15 min mit Katalysator) als mit fdC (Vollständiger Umsatz nach 4-6 h mit Katalysator). Es wird vermutet, dass dies an der stabilisierten Struktur von fdC liegt.^[225] Der Sauerstoff der Formyl-Gruppe befindet sich räumlich gesehen nahe an der C4 Amin-Gruppe und kann möglicherweise eine Wasserstoffbrückenbindung eingehen. Die fU Base besitzt im Gegensatz dazu keine Möglichkeit eine intramolekulare Wasserstoffbrücke auszubilden. Neben der Bindung von fdC bietet das Hydroxylamin **18** damit die zusätzliche Möglichkeit zwischen fdC, fdU und abasischen Stellen zu unterscheiden.

4.3.2. Ausblick

Mit den hier präsentierten Oligonukleotiden sollte nun, in Verbindung mit dem Hydroxylamin **18**, eine akkurate Aussage über die dynamische Symmetrie der DNA-Modifikation mdC und dessen oxidierten Derivate möglich sein. Durch zeitaufgelöste Untersuchungen sollte gezeigt werden können, nach welchen Kriterien die Modifikationen aufrechterhalten bzw. entfernt werden. So lassen sich dynamische epigenetische Muster auf dsDNA während der Zellalterung, der Zellteilung oder der Zelldifferenzierung untersuchen. Auch sollte damit untersucht werden können, ob die Oxidation eines mdCs des *sense* Strangs Auswirkungen auf den *antisense* Strang hat und *vice versa*.

Aktuell wird die Funktionsfähigkeit der hier erfolgreich synthetisierten *Hairpin*-Oligonukleotide, in Kombination mit dem Hydroxylamin **18**, in Sequenzierungsversuchen von *Pascal Giehr* von der *Arbeitsgruppe Epigenetik der Universität des Saarlandes (Prof. Dr. Jörn Walter)* getestet. Es konnte gezeigt werden, dass die Ligation der *Hairpin*-Oligonukleotide mit Schnittfragmenten der korrespondierenden Restriktionsenzyme sehr gut funktioniert. Erste vorläufige Sequenzierungsergebnisse weisen zudem darauf hin, dass mit dem Hydroxylamin **18** eine nahezu vollständige Abdeckung der fdCs erreicht werden konnte. Damit können diese nicht mehr durch Bisulfit deformyliert werden und erscheinen weiterhin als C im BS-Seq.

Neben der Sequenzierungsmethode mit dem Hydroxylamin **18** wurden noch weitere BS-Seq basierte Methoden, zur Sequenzierung von fdC auf ihre Qualität hin überprüft. Zum einen wurde das 2019 entwickelte ACE-Seq untersucht.^[226] Zum anderen wurde das MAB-Seq^[92, 93] und mit Bodipy^[227] ein weiteres Hydroxylamin getestet. Jedoch konnte keine dieser Methoden einen vergleichbaren Grad des Schutzes der fdCs erreichen bzw. fdC vollständig umsetzen. Abschließende Ergebnisse und Vergleiche werden in Kürze erwartet.

5. Wasserlösliches IBX zur Oxidation von hmdC-Oligonukleotiden

5.1. Wasserlösliches IBX-Derivat mit negativer Ladung

Alle bisher vorgestellten Methoden zur Sequenzierung von hmdC haben Limitierungen in ihrer Reaktivität, in der Reproduzierbarkeit der Daten oder der Einsetzbarkeit. Daher habe ich mich mit dem milden und selektiven Oxidationsmittel IBX, als weitere Alternative, beschäftigt.

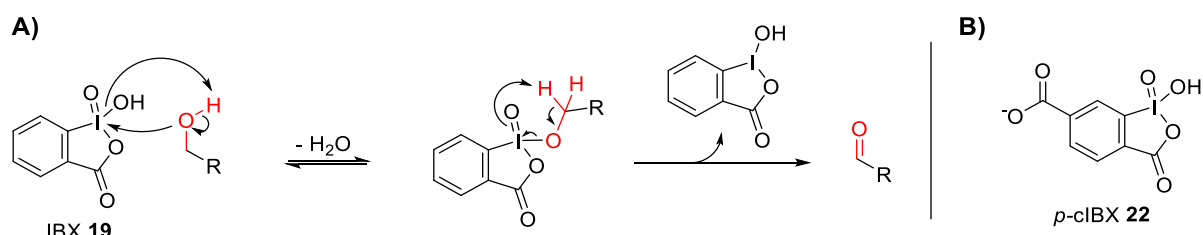


Abbildung 48: A) Mechanismus der Oxidation eines primären Alkohols mittels 2-Iodoxybenzoesäure (IBX) **19** zu dem entsprechenden Aldehyd. B) Struktur des negativ geladenen wasserlöslichen Analogons *p*-cIBX **22**.

Die hypervalente Iod Verbindung IBX **19** ist ein mildes und selektives Oxidationsmittel und kann primäre Alkohole zu Aldehyden oxidieren (Abbildung 48). Wasserlösliche IBX-Derivate finden immer mehr Anwendung in synthetischen Applikationen.^[228] Eine einfache Möglichkeit die Löslichkeit von IBX in Wasser zu ermöglichen, ist die Einführung einer Carboxylfunktion.^[229] Ein derartig modifiziertes IBX-Derivat könnte zur Oxidation von genomischer DNA eingesetzt werden. Die Wechselwirkungen zwischen der negativ geladenen Carboxylgruppe des IBX-Derivats und dem negativ geladenen Phosphatrückgrat der DNA sollten nicht stören, da dieses weit von der Hydroxylgruppe des hmdC entfernt ist. Möglicherweise kann die Carboxylgruppe, durch ihren leichten elektronenziehenden Effekt, sogar zusätzlich stabilisierend auf das pentavalente Iod wirken. Die in Abbildung 48 dargestellte *p*-Carboxyl-*o*-iodoxybenzoesäure (*p*-cIBX) **22** sollte daher im weiteren Verlauf auf ihr Potential, hmdC selektiv zu fdC zu oxidieren, untersucht werden.

5.1.1. Synthese des IBX-Derivates mit negativer Ladung (*p*-cIBX)

Die Synthese von *p*-cIBX **22** wurde beginnend mit 2-Aminoterephtalsäure **19** in zwei Schritten durchgeführt (Abbildung 49).^[229]

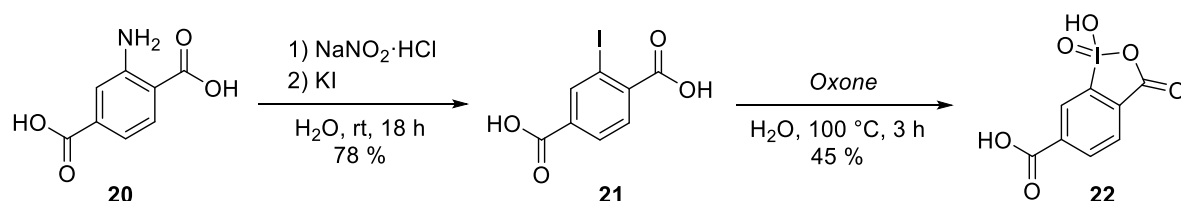


Abbildung 49: Syntheseroute zur Darstellung vom *p*-cIBX **22**, ausgehend von 2-Aminoterephtalsäure **19** in 2 Schritten.

Ausgehend von 2-Aminoterephthalsäure (**20**) wurde in einer Sandmeyer ähnlichen Reaktion 2-Iodoterephthalsäure (**21**) dargestellt. Anschließend wurde 2-Iodoterephthalsäure mit OXONE®, einer Mischung aus Kaliumperoxomonosulfaten, zum *p*-cIBX **22** oxidiert. **22** präzipitiert beim Abkühlen in den hier vorliegenden sauren Bedingungen und kann damit vom wasserlöslichen OXONE® abgetrennt werden.

5.1.2. Oxidation von hmdC mittels *p*-cIBX im Oligonukleotid-Kontext

Für die Oxidations-Experimente mit dem synthetisierten IBX-Derivat **22** wurde ein DNA-Oligonukleotid entworfen und synthetisiert, welches keine Sekundärstrukturen ausbildet und genau eine hmC-Base besitzt. Die Versuche wurden aufgrund der großen Anzahl, der durch das *screening* anfallenden Messungen, hauptsächlich über analytische RP-HPLC verfolgt.

Tabelle 8: Die Sequenz des Test-Oligonukleotids, dass in dieser Arbeit für die oxidationsversuche mit IBX-Derivaten verwendet wurde. Das Oligonukleotid besitzt eine hmC-Base und wird von kanonischen Basen umrahmt, so gewählt, dass möglichst keine Sekundär Strukturen ausgebildet werden.

ODN 17 | 5'-GTA ATG **hmd**CGC TAG G-3'

Um ein erkennbares und gut integrierbares Signal in der RP-HPLC zu erreichen, wurde die Sequenzlänge des Oligonukleotids auf 13 Basen limitiert (Tabelle 8). Damit konnte eine ausreichende Verschiebung der Retentionszeiten zwischen dem hmdC- und dem neu entstandenen fdC-Oligonukleotid erreicht werden. Die Verschiebung betrug dabei 0.5 min und konnte wie in Abbildung 50 ersichtlich klar erkannt und definiert werden.

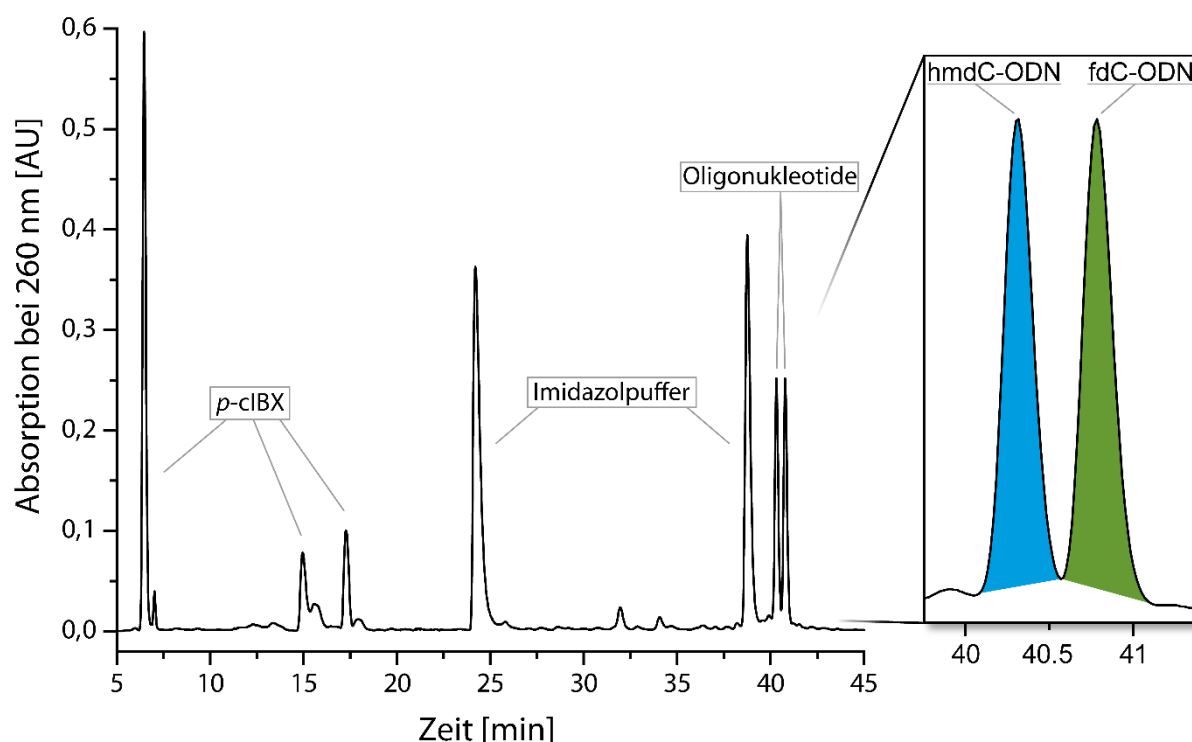


Abbildung 50: Beispiel eines HPLC Chromatogramms mit ca. 50 % Umsatz. Die ersten drei Signale bei 6 min, 15 min und 17 min sind die drei Oxidationsstufen des *p*-cIBX. Die Signale bei 24-25 min und bei 38 min sind dem Imidazol-Puffer zuzuschreiben. Bei 40-40.5 min ist das hmdC-Oligonukleotid Signal (blau) zu sehen. Bei 40.5-41 min ist das neu gebildete fdC-Oligonukleotid Signal (grün) zu erkennen.

Bei den finalen Experimenten wurden nach erfolgreicher Oxidation, die Oligonukleotide zusätzlich mittels EtOH-Fällung aus der Reaktionslösung präzipitiert und die Menge von fdC im Vergleich zu hmdC mittels LC-MS/MS bestimmt. Dabei wurde auch die Menge von cadC bestimmt, um eine weitere Oxidation zur Carbonsäure ausschließen zu können. Die LC-MS/MS Messungen wurden zusammen mit Sarah Schiffers durchgeführt.

Ein typisches Experiment wurde in 80 µL Gesamtvolumen durchgeführt (Tabelle 9). Dieses Volumen wurde gewählt, um bei der Injektion in die RP-HPLC Spielraum hin zum maximalen Injektionsvolumen von 100 µL zu haben und gleichzeitig möglichst wenig Material bei der Injektion zu verlieren. Für die Experimente wurde im Folgenden jeweils 300-400 pmol des synthetisierten ORN 7 verwendet. Mit dieser Menge konnten die Absorptions-Signale der HPLC zuverlässig zugeordnet werden und vom Rauschen bzw. von Säulenverunreinigungen unterschieden werden. Als Puffer wurde ein Imidazol-Puffer in unterschiedlichen Konzentrationen und pH-Werten im basischen Bereich verwendet, da sich die oben beschriebenen IBX-Derivate nur schlecht in sauren Bedingungen lösen. Es wurden auch weitere Puffersysteme untersucht (Na₂H₂PO₄/Na₂HPO₄; Diglycin Puffer (GlyGly); 2-(Cyclohexylamino)-ethansulfonsäure Puffer (CHES)), jedoch führten diese zu einer signifikanten Verschlechterung, bis hin zum Ausbleiben der Oxidation. Gängige Puffersysteme wie HEPES oder TRIS schieden aus, da sie Hydroxygruppen besitzen, welche oxidiert werden könnten und damit die IBX-Derivate abfangen würden.

Tabelle 9: Typischer Reaktionsansatz zur Untersuchung der Oxidation von hmdC- zu fdC-Oligonukleotiden mittels des negativ geladenen IBX-Derivats *p*-cIBX 22.

Bezeichnung	Volumen [µL]	Endkonzentration
ODN 17	1.5-2	300-400 pmol in 80 µL
Imidazol-Puffer (1M, pH 8-9)	60	0.375 -1.5 M
IBX-Derivat	-	50 – 10000 Äquivalente
ddH ₂ O	18-18.5	-

Da sich zeigte, dass die Oxidation mit Hilfe des IBX-Derivats 22 bei Raumtemperatur keinen messbaren Umsatz zeigte, wurde die Reaktion im Weiteren bei einheitlich 60 °C durchgeführt. Auf diese Weise konnte eine Reaktion, bei zugleich geringer Degradation der eingesetzten DNA, erreicht werden.

Im ersten Schritt wurde der Umsatz von hmdC zu fdC in Abhängigkeit des pH-Werts untersucht, bei pH-Werten von 7 bis 9 (Abbildung 51). Der obere pH-Bereich wurde gewählt, da durch die zunehmende Protonierung der Carbonsäure unterhalb eines pH-Werts von 7 die Wasserlöslichkeit des IBX-Derivats rasch abnimmt. Die obere Grenze des pH-Werts bei 9 wurde aufgrund der pK_a Werte der Basen Guanin (9.2) und Thymin (9.7) beschränkt.^[230] Bei überwiegend negativ geladenen Basen könnte das auch negativ geladene IBX-Derivat aufgrund von ionischer Abstoßung nicht mehr mit dem Ziel-Oligonukleotid wechselwirken.

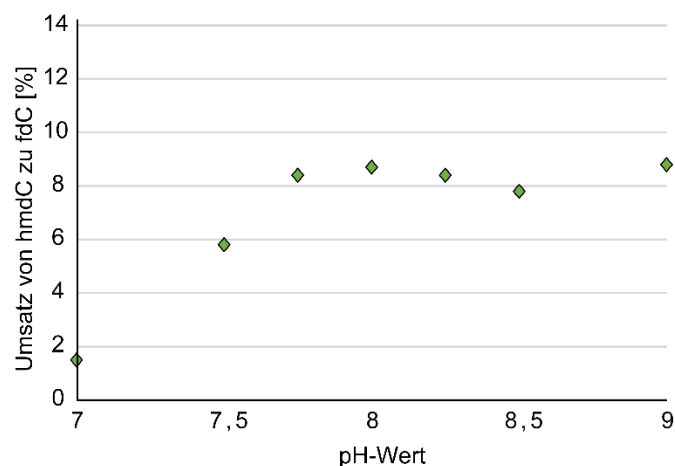


Abbildung 51: Umsatz vom hmdC zum fdC Oligonukleotid mit 100 Äq *p*-cIBX, bei 60 °C für 4 h, mit pH-Werten von 7-9 (0.75 M Imidazol-Puffer). Die grünen Rauten zeigen den Umsatz in Prozent bei den jeweiligen pH-Werten. Die rot gestrichelte Linie zeigt eine Trendlinie der pH-Abhängigkeit des Umsatzes.

Wie aus Abbildung 51 ersichtlich ist, steigt die Reaktivität von *p*-cIBX nahezu proportional zum pH Wert an und findet ein Plateau bei pH 8-9. Die Reaktionen wurden daher im Weiteren in Imidazol-Puffern mit pH 9 durchgeführt. Die Konzentration des Puffers wurde aufgrund des leicht außerhalb der Maximalen Pufferleistung von Imidazol liegende pH Werts auf 1.5 M erhöht. Damit konnten auch höhere Konzentrationen an IBX-Derivat **22** in den 60 µL Puffer gelöst werden.

Im Weiteren wurde der Zusammenhang zwischen unterschiedlichen Überschüssen an *p*-cIBX und dem Fortschreiten der Reaktion untersucht (Abbildung 52). Wie zu erwarten, konnte gezeigt werden, dass bei einem höherer Überschuss an *p*-cIBX auch die Reaktion schneller abläuft.

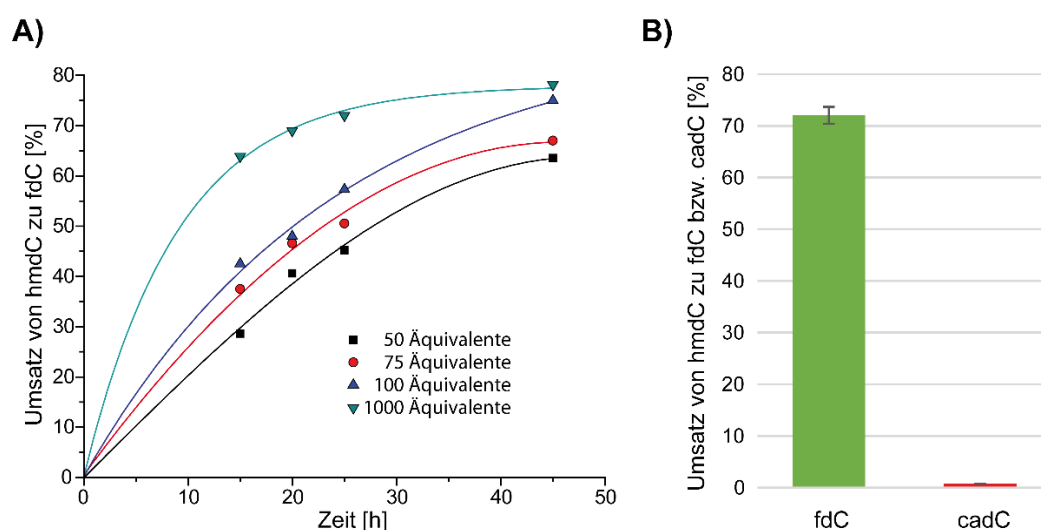


Abbildung 52: A) Umsatz vom hmdC zum fdC Oligonukleotid bei 60 °C und mit Imidazol-Puffer (1.5 M, pH 9) für die Zeitpunkte 15, 20, 25 und 45 h. Der Überschuss an *p*-IBX betrug 50 (Quadrat), 75 (Kreis), 100 (blaues Dreieck) und 1000 (türkises Dreieck) Äquivalente in Bezug auf hmdC. Aufgrund von Übersichtlichkeit wurden Trendlinien gezogen. B) Menge an fdC und cadC im Vergleich zu hmdC ermittelt mittels quantitativer LC-MS/MS Messung (pH 9, 1000 Äq, 22 h und 60 °C).

Dies wird vor allem im großen Sprung zwischen 100 Äq und 1000 Äq ersichtlich. Trotz des 10 fachen Überschusses, zeigen beide Reaktionsansätze bei einer Reaktionszeit von 45 h mit einem Maximum von 70-80 % Umsatz kaum Unterschiede. Die quantitativen LC-MS/MS Untersuchungen von *Sarah Schiffrs* konnten den Umsatz mit 1000 Äq, welcher durch das Integrieren der RP-HPLC Chromatogramme ermittelt wurde, bestätigen. Zudem zeigten die Daten, dass das IBX-Derivat hmdC selektiv zu fdC oxidiert wurde. Die höhere Oxidationsstufe cadC wurde nur mit sehr geringen Mengen gemessen (0.7 %). Diese Eigenschaft ist beispielsweise für eine Anwendung in der CLEVER-Seq essentiell. Eine Plateau Bildung bei 70-80 % des Umsatzes wurde nicht erwartet, da der hohe Überschuss das Gleichgewicht der Reaktion theoretisch in Richtung fdC schieben sollte. Daher wurde ein weiteres Experiment durchgeführt, in dem der Überschuss auf 5000 bzw. 10000 Äquivalente erhöht wurde, um das Gleichgewicht noch stärker in Richtung fdC zu schieben. In der Theorie sollte hier das Maximum bei 70-80 % erreicht bzw. übertroffen werden. Interessanterweise fällt aber, wie aus Abbildung 53 A ersichtlich, die Umsetzung von hmdC zu fdC bei steigenden Äquivalenten (5000 und 10000) von *p*-cIBX dramatisch ab.

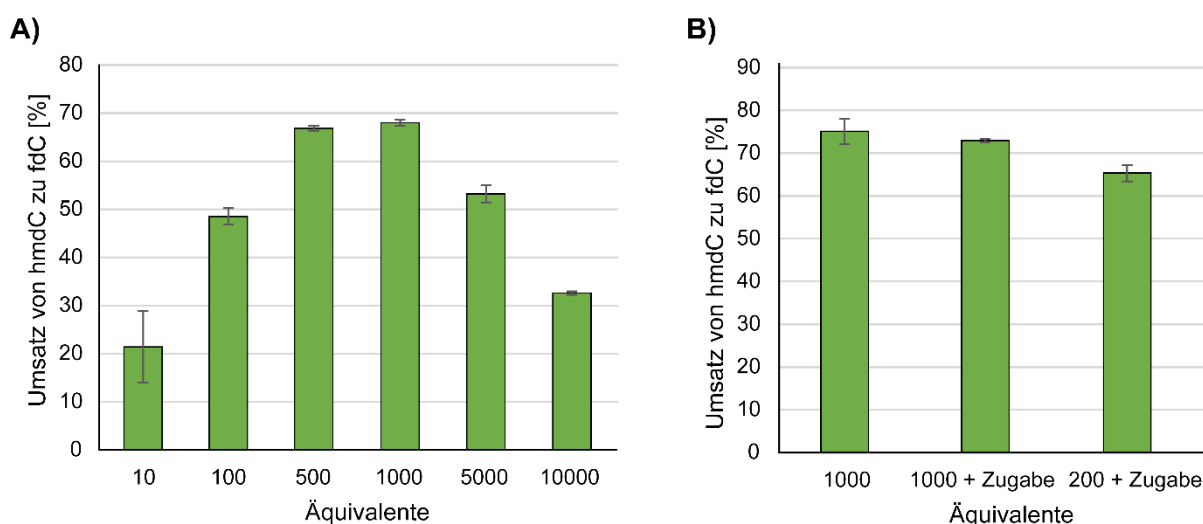


Abbildung 53: A) Umsatz von hmdC zu fdC mit 10, 100, 500 1000, 5000 und 10000 Äquivalenten *p*-cIBX, bei 60 °C für 20.5 h, mit Imidazol-Puffer (1.5 M, pH 9) zeigt ein Maximum bei 500-1000 Äquivalenten. B) Umsatz von hmdC zu fdC mit *p*-cIBX, bei 60 °C für 30 h, mit Imidazol-Puffer (1.5 M, pH 9). Mit 1000 Äquivalenten, mit 1000 Äquivalenten und erneuter Zugabe von 200 Äquivalenten bei 5 h und 20 h bzw. 200 Äquivalenten und erneuter Zugabe von 200 Äquivalenten bei 5 h und 20 h.

Die höchste Umsatzungsrate wird weiterhin bei 1000 Äquivalenten erreicht. In einem weiteren Versuch wurden zu Reaktionen mit 1000 bzw. 200 Äquivalenten nach 5 h bzw. 20 h jeweils 200 weitere Äquivalente des IBX-Derivats zugegeben (Abbildung 53 B). Damit sollte ein Zusammenhang mit einer möglichen Zersetzung des *p*-cIBX unter den Reaktionsbedingungen ausgeschlossen werden. Auch diese zeitlich versetzte Zugabe von *p*-cIBX brachte keine erhöhten Umsatzwerte. Eine mögliche Erklärung für das beschriebene Reaktionsverhalten, könnte in einer konzentrationsabhängigen Verschlechterung der Löslichkeit bzw. Reaktivität von *p*-cIBX liegen. IBX besitzt eine polymer Struktur, welche durch intermolekulare sekundäre Bindungen zwischen Iod (V) und Sauerstoff gebildet

werden.^[231] Möglicherweise ist die Tendenz zur Bildung dieser Polymerstruktur beim *p*-cIBX mit der Konzentration in der Lösung verknüpft. Bei einer solchen Clusterbildung würde das innen liegende IBX-Derivat nicht mehr für eine Reaktion zur Verfügung stehen.

Durch diese Arbeit konnte aber gezeigt werden, dass sich wasserlösliches IBX zur Oxidation von hmdC zu fdC in Oligonukleotiden prinzipiell eignet. Die Umsatzrate von 70-80 % ist dabei durchaus vielversprechend. Bei einer anschließenden Anreicherung der fdC haltigen DNA, nach dem Vorbild von *Matsushita et al.*,^[176] könnten damit erste Sequenzierungsversuche unternommen werden. Jedoch müsste dafür die relativ lange Reaktionszeit (20-45 h) weiter optimiert werden.

5.2. Wasserlösliches IBX-Derivat mit positiver Ladung

Nachteil der durch die Carboxyl-Gruppe erreichten Wasser-Löslichkeit, ist die Limitierung auf Puffersysteme mit einem basischen pH-Wert. Daher sollte die Löslichkeit stattdessen mit einer permanent positiven Ladung, einem quartären Amin, erreicht werden. Zum einen hat dies den Vorteil, dass das Molekül unabhängig vom Puffer-Milieu als Salz vorliegt. Damit bleibt die Löslichkeit in H₂O auch in Puffersystemen unterschiedlicher pH-Werte erhalten. Zum anderen könnte eine ionische Bindung an das, bei basischem pH-Wert, negativ geladene DNA Rückgrat von großem Vorteil sein. Durch diese Bindung wäre das IBX Analog an der DNA gut positioniert um eine erfolgreiche Oxidation durchzuführen. Um diese aktive Positionierung nutzen zu können muss der funktionelle Bereich des IBX bis hin zur hmC-Base reichen können. Das heißt, die IBX Funktion muss in einem Abstand von der positiven Ladung platziert werden, die mindestens so groß ist wie der Abstand des DNA-Rückgrats zum primären Alkohol der hmC-Base.

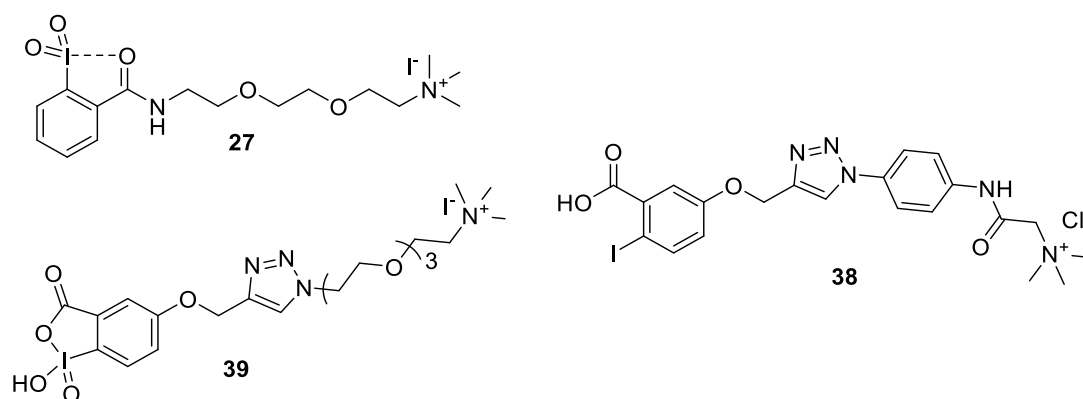


Abbildung 54: Übersicht der in dieser Arbeit synthetisierten positiv geladenen und wasserlöslichen IBX-Derivate IBX-Amid **27** und PEG⁺-IBX **39** und die Vorstufe SF⁺-IBX **34**.

Dieser Abstand wurde zum einen durch eine mono Peptidkopplung an ein Diamin (IBX-Amid **27**) erreicht (Abbildung 54). Amid-Derivate von IBX besitzen analog zu IBX ein pentavalentes Iod. Das pentavalente Iod kann mit dem Sauerstoff der Peptidbindung, eine intramolekulare Bindungen eingehen. Die pseudocyclische Struktur der IBX-Amide erhöht ihre Löslichkeit und besitzt oxidative Eigenschaften, die mit IBX vergleichbar sind. Die Reaktionsbedingungen sind mit Raumtemperatur sehr mild und daher kompatibel mit DNA.^[232] Möglicherweise zeigt das IBX-Amid, im Gegensatz zu *p*-cIBX, eine bessere Performance bei der Erhöhung der Äquivalente bzw. Konzentration.

Zum anderen wurde IBX, mittels kupferkatalysierten 1,3-dipolaren Cycloaddition (CuAAC), mit zwei unterschiedlichen, positiv geladenen Resten verbunden. Die permanent positive Ladung wurde zuerst über den Azid-Baustein **32** eingefügt, der auch zur Synthese des in der Arbeitsgruppe *Carell* etablierten Hydroxylamins „*SuperFly*“ (SF) verwendet wird.^[223] Die Bindung des Hydroxylamins an fdC konnte in dieser Arbeit bereits gezeigt werden (Abbildung 47). Daher ist es naheliegend, dass ein SF⁺-IBX-Derivat **34**, bei ähnlicher Größe, auch an hmdC heran reichen kann. Ein mögliches π -stacking

mit dem Aromaten des IBX-Rests und dem Aromaten des SF-Bausteins könnte die Positionierung positiv beeinflussen und dadurch zu einer Reaktivitätssteigerung führen. Als zweites wurde ein PEG-Linker **36** gewählt. Durch die Flexibilität des Polyethylenglycols ist die Position des IBX nicht strikt vorgegeben und kann daher das hmdC leicht erreichen. Auch hier kann der Aromat des IBX-Rests die Positionierung, durch ein mögliches π -stacking, unterstützen.

5.2.1. Synthese der IBX-Derivate mit positiver Ladung

5.2.1.1. Synthese des IBX-Amids

Das IBX-Amid **27** wurde ausgehend von dem Säurechlorid **23** über 3 Schritte synthetisiert (Abbildung 55).

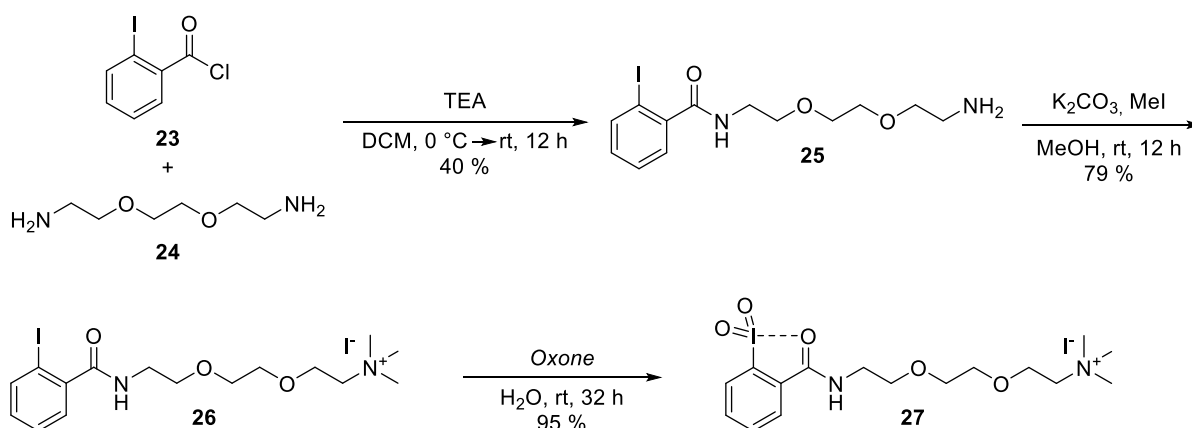


Abbildung 55: Syntheseroute zur Herstellung vom IBX-Amid **27** ausgehend vom Säurechlorid **23** und dem Diamin **24**.

Zunächst wurde das Säurechlorid **23** mit dem Diamin **24** zum monosubstituierten Produkt **25** umgesetzt.^[233] Hierbei wurde das Diamin **24** nur in geringem Überschuss eingesetzt, da Amine in der Säulenchromatographie zu sehr großen Elutionsbereichen neigen und damit nur schwer vom Produkt zu trennen sind. Eine vorhergehende teilweise Schützung des Aminrestes mit *tert*-Butyloxycarbonyl wurde versucht, führte aber über drei Schritte zu geringeren Ausbeuten. Die Methylierung von **25** wurde mit Iodmethan in guten Ausbeuten erreicht. Bemerkenswert war, dass das positiv geladene Produkt **26** mittels Normalphasen Säulenchromatographie gereinigt werden konnte. Die erfolgreich synthetisierte IBX-Amid-Vorstufe **26** konnte für mehrere Wochen bei Raumtemperatur gelagert werden.

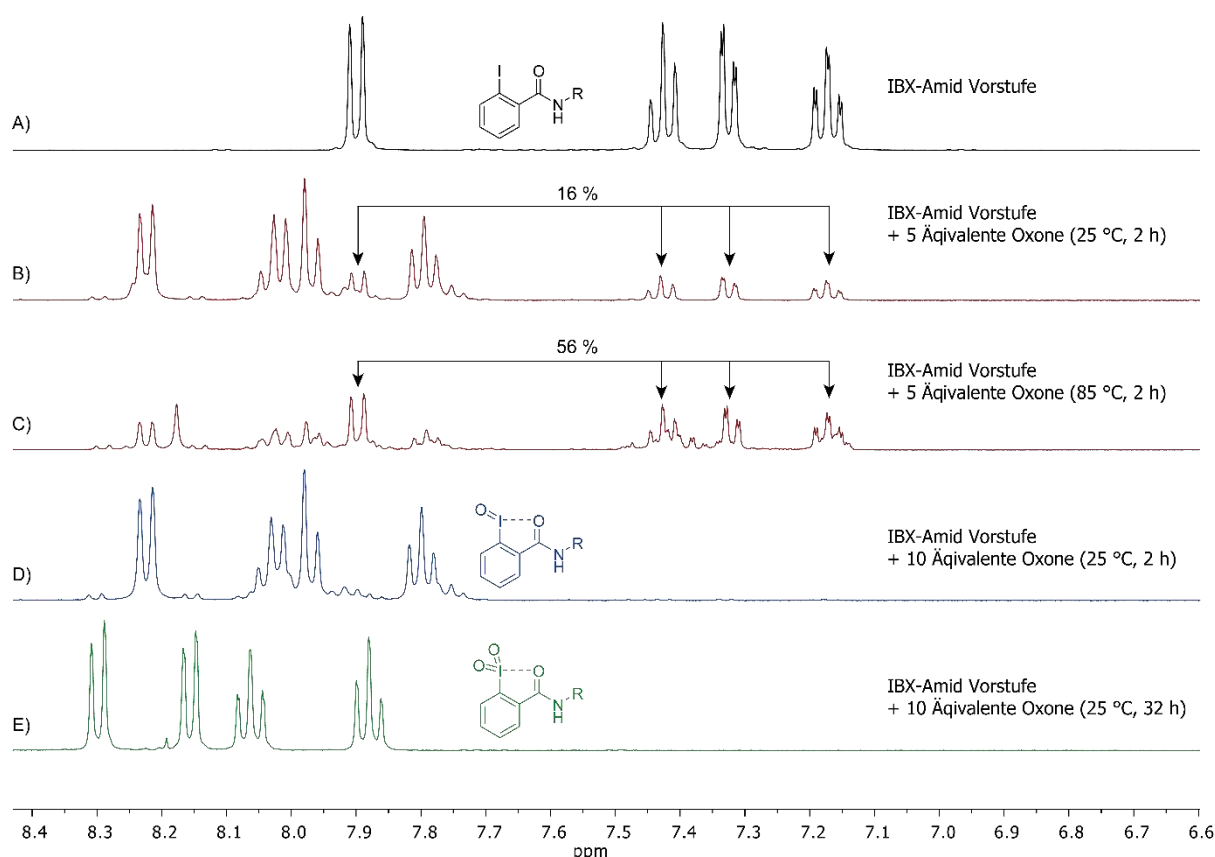


Abbildung 56: Ausschnitt der NMR Signale von **26** im aromatischen Bereich, bei unterschiedlichen Oxidationsbedingungen. A) IBX-Amid-Vorstufe **26**, ohne Oxidationsmittel. B) IBX-Amid-Vorstufe **26** umgesetzt mit 5 Äq OXONE® bei 25 °C für 2 h. Die schwarzen Pfeile markieren die Signale der IBX-Amid-Vorstufe **26** (16 % im Vergleich zu neu gebildetem Material). C) IBX-Amid-Vorstufe **26** umgesetzt mit 5 Äq OXONE® bei 85 °C für 2 h. Die schwarzen Pfeile markieren die Signale der IBX-Amid-Vorstufe **26** (56 % im Vergleich zu neu gebildetem Material). D) IBX-Amid-Vorstufe **26** umgesetzt mit 10 Äq OXONE® bei 25 °C für 2 h (Blau, erste Oxidationsstufe). E) IBX-Amid-Vorstufe **26** umgesetzt mit 10 Äq OXONE® bei 25 °C für 32 h (Grün, zweite Oxidationsstufe).

Die Oxidation von **26** zum IBX-Amid **27** wurde zunächst bei 73 °C mit drei Äquivalenten OXONE® durchgeführt,^[233] dies führte jedoch zu keiner Oxidation. Auch der Ersatz von OXONE® mit KBrO₃ führte zu keiner Umsetzung. Um gute Bedingungen für die Oxidation zu finden, wurde die Umsetzung mit unterschiedlicher Dauer, Temperaturen und Äquivalenten, mittels NMR-Messungen untersucht. Um Zeit und Edukte zu sparen, wurden die Testreaktionen direkt in deuteriertem Wasser durchgeführt. Eine Erhöhung der Temperatur auf 85 °C und auf 5 bzw. 10 Äquivalente OXONE® führten statt zum IBX-Derivat, zur Zersetzung der IBX-Amid-Vorstufe **26** (Abbildung 56 C). Daher wurde die Temperatur drastisch auf 25 °C verringert, da auch bei 40 °C eine Zersetzung beobachtet werden konnte. Es zeigte sich, dass bei Raumtemperatur nach 2 h Reaktionszeit die erste Oxidationsstufe, ein Trivalentes Iod, zu 84 % (5 Äq, Abbildung 56 B) bzw. vollständig (10 Äq, Abbildung 56 D) erreicht wurde. Signale von Nebenprodukten entstanden bei 25 °C nicht. Die Reaktion mit 10 Äquivalenten wurde daher für 32 h bei 25 °C durchgeführt und ausgehend vom NMR Spektrum und Massenspektrometrie, eine vollständige Umwandlung zur erwünschten Oxidationsstufe erreicht.

Anders als bei *p*-cIBX **22** konnte das IBX-Amid **27** nicht aus der wässrigen OXONE® Lösung gefällt werden. Stattdessen wurde das OXONE® durch die Zugabe von MeCN/H₂O (2/1) gefällt.^[234] Das Produkt war bei 25 °C stabil, in Lösung bei 60 °C jedoch konnte nach weniger Stunden Zersetzung im NMR beobachtet werden.

5.2.1.2. Synthese der Alkin-IBX Vorstufe

Um die beiden über CuAAC zugänglichen, positiv geladenen und wasserlöslichen IBX Analoga darzustellen, musste zunächst die ethylgeschützte IBX Vorstufe mit einer Alkingruppe **31** synthetisiert werden (Abbildung 57).

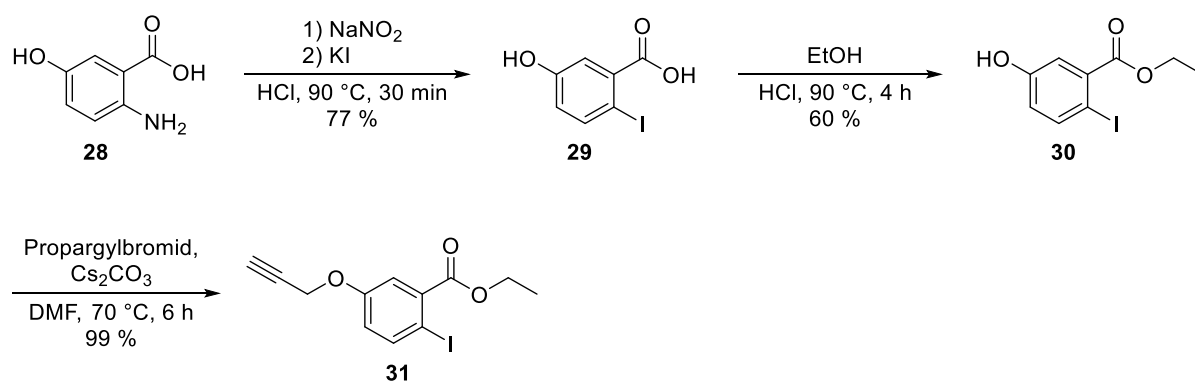


Abbildung 57: Syntheseroute zur Darstellung der Alkin-IBX Vorstufe **31**, ausgehend von 2-Amin-5-hydroxybenzoesäure (**28**).

Die Amin-Gruppe von 2-Amin-5-hydroxybenzoesäure (**28**) wurde mittels Sandmeyer ähnlicher Reaktion mit Iod substituiert. Da der direkte Umsatz von **29** mit Propargylbromid kein gewünschtes Produkt lieferte, wurde die Carbonsäure in einer Fischer-Veresterung mit Ethanol in konzentrierter HCl zum 2-Iod-5-hydroxybenzoesäureethylester (**30**) verestert.^[235] Im folgendem Schritt wurde dann die Hydroxygruppe von **30** mit Cs₂CO₃ deprotoniert und mit Propargylbromid zur Alkin-IBX Vorstufe **31** umgesetzt.

5.2.1.3. Synthese des SF⁺-IBX-Derivats

Als nächstes wurde der Baustein **32** des „SuperFly“ mit permanent positiv geladenem quartärem Amin synthetisiert. Die Synthese ist literaturbekannt und wurde in vergleichbaren Ausbeuten erreicht.^[223]

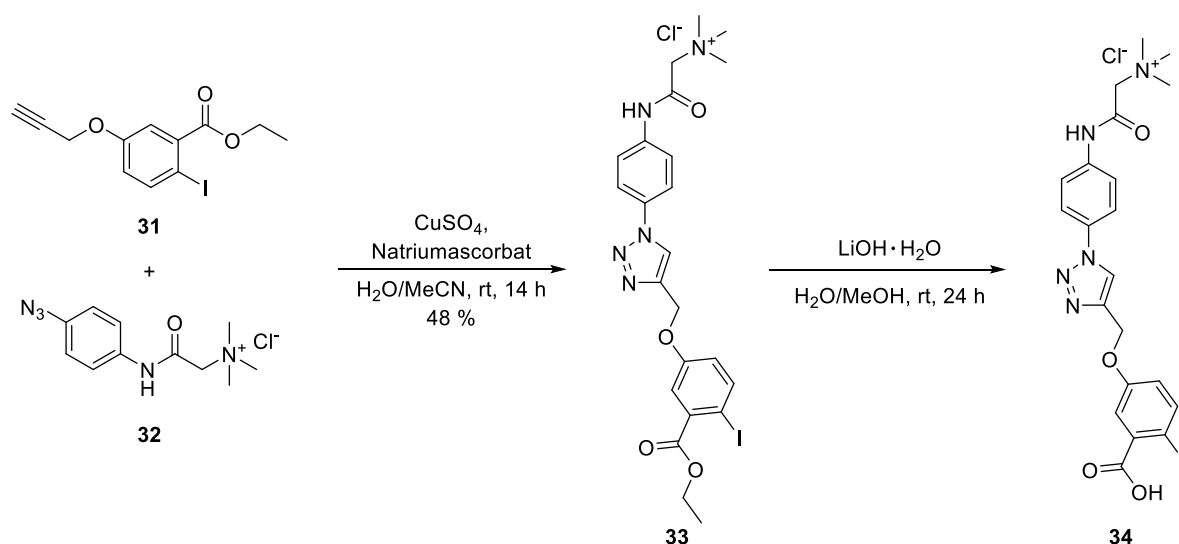


Abbildung 58: CuAAC und drauf folgende Lithiumhydroxid Entschützung zur Darstellung der Vorstufe **34** des SF^+ -IBX.

Bei der CuAAC von **32** mit der geschützten Alkin-IBX Vorstufe **31** konnte eine Präzipitation beobachtet werden. Entgegen der Erwartungen konnten weder die geschützte **33** noch die entschützte Vorstufe **34** des SF^+ -IBX in Wasser gelöst werden. Das Molekül konnte nur in DMSO vollständig in Lösung gebracht werden. Da die Wasserlöslichkeit die Grundvoraussetzung für eine Interaktion mit DNA darstellt, wurde das SF^+ -IBX im Weiteren nicht mehr verwendet.

5.2.1.4. Synthese des PEG^+ -IBX-Derivats

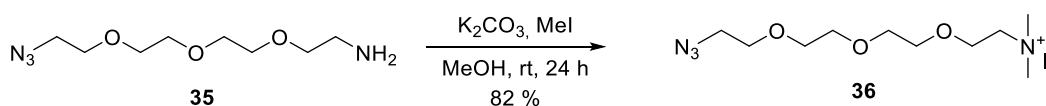


Abbildung 59: Reaktionsschritt zur Darstellung des PEG-Linkers **36**, ausgehend von 11-Azid-3,6,9-trioxaundecan-1-amin (**35**).

Das kommerziell erhältliche 11-Azid-3,6,9-trioxaundecan-1-amin (**35**) wurde mit Iodmethan in guter Ausbeute bis zum quartären Amin **36** methyliert. Durch Lösen in DCM konnte der PEG-Linker **36** von den wasserlöslichen Salzen getrennt werden.

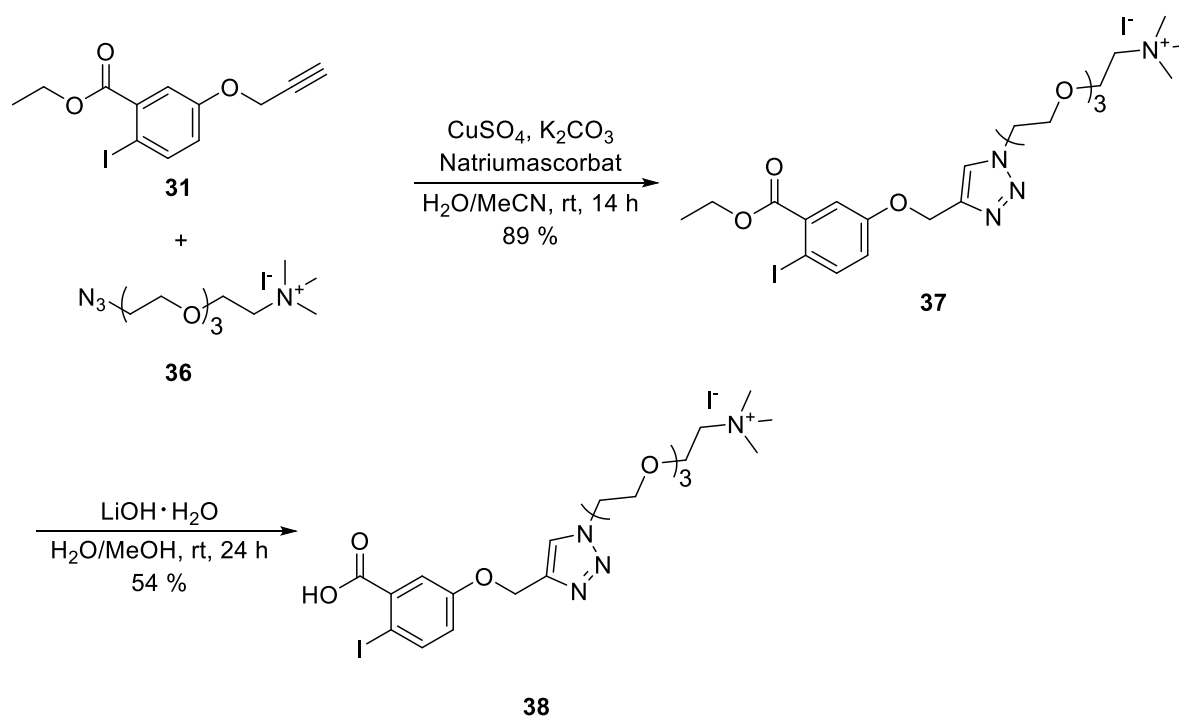


Abbildung 60: CuAAC und drauf folgende Lithiumhydroxid Entschützung zur Darstellung der Vorstufe **38** des PEG^+ -IBX.

Der PEG-Linker **36** wurde anschließend mit der geschützten IBX Vorstufe **31** in einer CuAAC verknüpft. Hierbei war interessanterweise ein basisches Milieu notwendig. Die Cycloaddition ohne K_2CO_3 brachte eine sehr viel geringere Ausbeute und bei weiterer Zugabe von Natriumascorbat und CuSO_4 , über mehrere Tage, zersetzten sich die Edukte. Daher wurde **36** nicht mehr, wie vorher beschrieben, vom K_2CO_3 Salz der Methylierung gereinigt, sondern in einer 1:1 (Gew. %) Mischung verwendet. Nach der Entschützung der Carboxylgruppe von **37** durch Lithiumhydroxid konnte die Vorstufe des PEG^+ -IBX **38** erhalten werden (Abbildung 60).

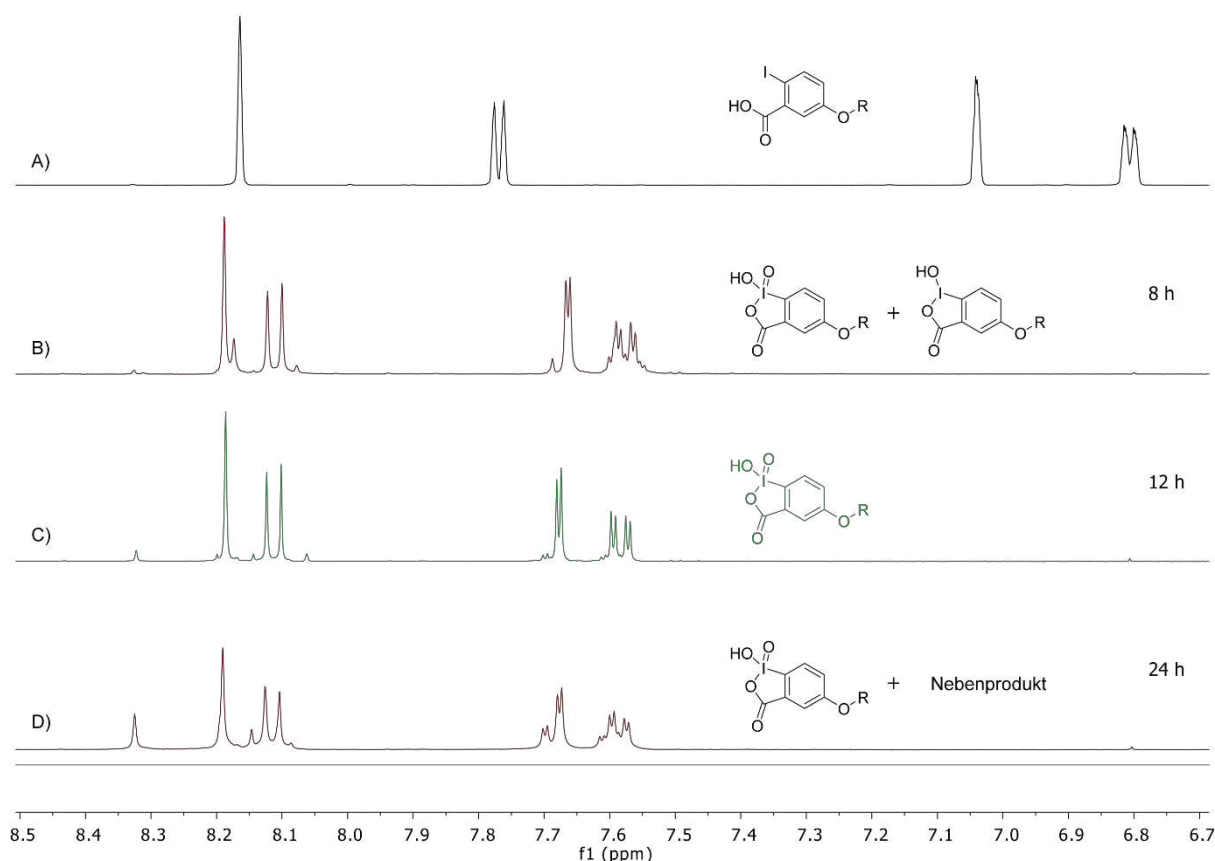


Abbildung 61: Ausschnitt der NMR Signale von **38** im aromatischen Bereich, bei unterschiedlichen Oxidationsbedingungen. A) PEG⁺-IBX Vorstufe **38**, ohne Oxidationsmittel. B) PEG⁺-IBX Vorstufe **38** umgesetzt mit 10 Äq OXONE[®] bei 25 °C für 8 h. C) PEG⁺-IBX Vorstufe **38** umgesetzt mit 10 Äq OXONE[®] bei 25 °C für 12 h. D) PEG⁺-IBX Vorstufe **38** umgesetzt mit 10 Äq OXONE[®] bei 25 °C für 24 h.

Auch bei der Oxidation der PEG⁺-IBX Vorstufe **38** führten hohe Temperaturen zu Nebenprodukten und zur Zersetzung. Daher wurde auch diese Reaktion bei 25 °C durchgeführt. Interessanterweise war **38** schon nach 12 h bei 10 Äquivalenten OXONE[®] vollständig zum PEG⁺-IBX **39** oxidiert. Eine längere Reaktionszeit führte zur Entstehung eines Nebenprodukts mit NMR-Signalen bei 8.33, 8.15 und 7.70 ppm (Abbildung 61).

5.2.2. Oxidation von hmdC mittels PEG⁺-IBX und IBX-Amid im Oligonukleotid-Kontext

Die Oxidationsexperimente wurden unter denselben Versuchsbedingungen, wie für *p*-cIBX beschrieben, durchgeführt (Tabelle 9). Auch hier wurde das Oligonukleotid ODN **17** (Tabelle 8) mit 13 Basen und genau einem hmdC verwendet. Das Probenvolumen betrug 80 µL und die Menge des Oligonukleotids 400 pmol. Wie schon bei der Synthese beschrieben, zersetzten sich beide Moleküle (PEG⁺-IBX **39** und IBX-Amid **27**) bei 60 °C. Daher wurden die ersten Experimente in 25 °C und 40 °C durchgeführt, um die IBX-Derivate möglichst lange stabil zu halten. Es wurden 100 und 500 Äquivalente als Startpunkt gewählt und ein Imidazol-Puffer (0.75 M) bei pH 8.0 und 8.5 verwendet.

Jedoch zeigten sich bei den verwendeten Bedingungen keine nennenswerten Umsätze. Die Bedingungen und dazugehörigen Reaktionsumsätze sind in Tabelle 10 zusammengefasst.

Tabelle 10: Verwendete Bedingungen für die Oxidation von hmdC zu fdC-Oligonukleotiden mit den IBX-Derivaten PEG⁺-IBX **39** und IBX-Amid **27**. Die Pufferkonzentration befand sich bei 0.75 M und die Reaktionszeit betrug 24 h. Die beiden besten Ergebnisse wurden *Fett* markiert.

A) PEG⁺-IBX:

Äquivalente	Temperatur [°C]	Puffer [pH]	Umsatz [%]
500	25	8.5	1.8
		8	0.0
	40	8.5	0.8
		8	2.7
100	25	8.5	0.0
		8	0.0
	40	8.5	0.1
		8	0.5

B) IBX-Amid:

Äquivalente	Temperatur [°C]	Puffer [pH]	Umsatz [%]
500	25	8.5	1.0
		8	0.9
	40	8.5	2.9
		8	1.0
100	25	8.5	0.8
		8	0.7
	40	8.5	1.7
		8	1.7

Da sich bei den Experimenten ein leichter Trend hin zu der höheren Temperatur 40 °C zeigte, wurde trotz möglicher Zersetzung der IBX-Derivate die Temperatur auf 60 °C erhöht. Dabei fiel auf, dass schon bei 400 Äquivalenten eine Zersetzung des verwendeten Oligonukleotids einsetzte. Versuche mit 1000, 5000 und 10 000 Äq konnten weder über HPLC noch über LC-MS/MS eindeutig ausgewertet werden. Daher wurde eine weitere Reaktionsreihe beginnend mit zwei Äquivalenten angesetzt (Abbildung 62).

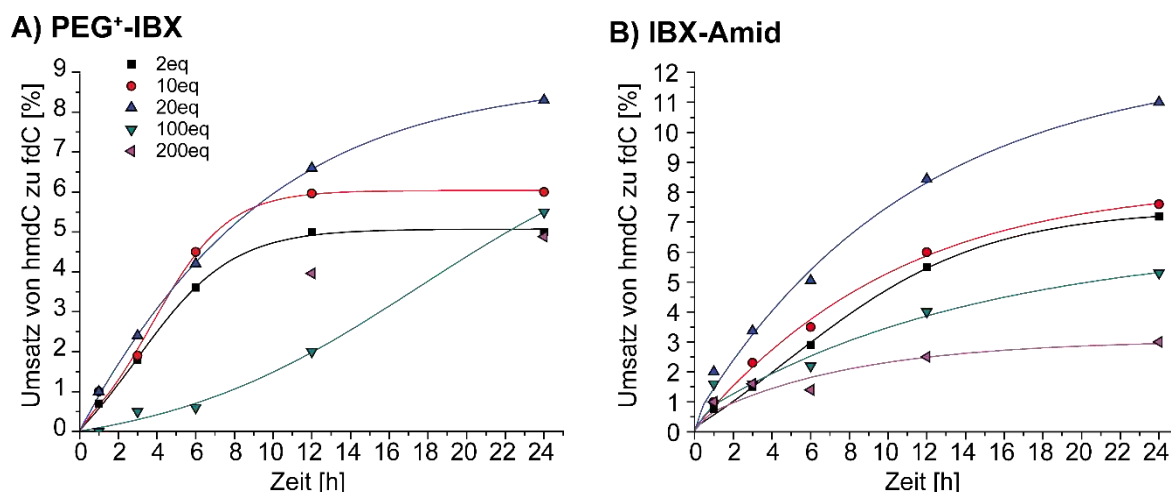
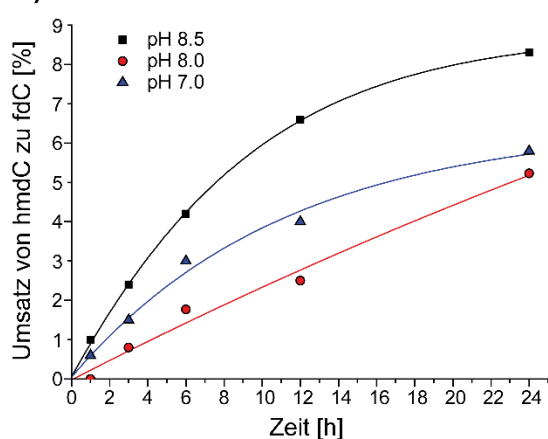


Abbildung 62: A) Umsatz vom hmdC- zum fdC-Oligonukleotid bei 60 °C und pH 8.5, mit 2 (schwarzes Viereck), 10 (roter Kreis), 20 (blaues Dreieck), 100 (grünes Dreieck), 200 Äq (rosa Dreieck) von PEG⁺-IBX **39** im zeitlichen Verlauf. B) Umsatz vom hmdC- zum fdC-Oligonukleotid bei 60 °C und pH 8.5, mit 2 (schwarzes Viereck), 10 (roter Kreis), 20 (blaues Dreieck), 100 (grünes Dreieck), 200 Äq (rosa Dreieck) von IBX-Amid **27** im zeitlichen Verlauf. Aufgrund einer besseren Übersichtlichkeit wurden jeweils Trendlinien gezogen.

Durch die Erhöhung der Temperatur auf 60 °C konnte nun eine merkliche Umsatzsteigerung im Vergleich zu 25 °C und 40 °C erreicht werden. Der Umsatz stieg bei beiden Derivaten schnell an und flachte hin zu 24 h langsam ab. Bei beiden positiv geladenen IBX-Derivaten wurde schon bei 2 Äquivalenten ein Umsatz von 5.1 % (PEG⁺-IBX **39**) und 7.2 % (IBX-Amid **27**) nach 24 h Reaktionszeit erreicht. Der Umsatz steigerte sich bis zu 20 Äq und erreichte hier sein Maximum von 8.8 % (PEG⁺-IBX **39**) und 11.0 % (IBX-Amid **27**). Wie bei den Beobachtungen von *p*-cIBX (Abbildung 53), wurde bei den Reaktionen ein Äquivalentenmaximum erreicht. Eine Erhöhung der Äquivalente auf 100 bzw. 200 waren in Hinsicht auf eine Umsatzsteigerung kontraproduktiv.

Da die Löslichkeit der IBX-Derivate auch im neutralen und sauren pH-Werten gewährleistet sein sollte, wurden im Weiteren unterschiedliche pH Werte untersucht, bei den sonst gleichbleibenden Bedingungen mit 20 Äq vom eingesetzten Reagenz.

A) PEG⁺-IBX

B) IBX-Amid

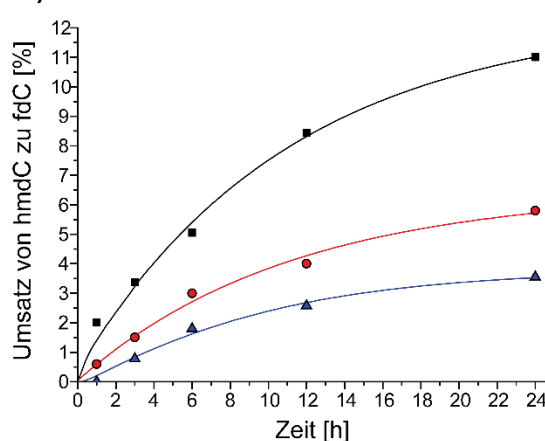


Abbildung 63: A) Umsatz von hmdC zu fdC bei 60 °C und 20 Äquivalenten von PEG⁺-IBX **39**, mit den pH-Werten 8.5 (schwarzes Viereck), 8.0 (blaues Dreieck), 7.0 (roter Kreis) im zeitlichen Verlauf. B) Umsatz von hmdC zu fdC bei 60 °C und 20 Äquivalenten von IBX-Amid **27**, mit den pH-Werten 8.5 (schwarzes Viereck), 8.0 (blaues Dreieck), 7.0 (roter Kreis) im zeitlichen Verlauf. Der pH-Wert für 5.0 konnte nicht aufgetragen werden, da sich das Oligonukleotid zersetzte. Aufgrund einer besseren Übersichtlichkeit wurden jeweils Trendlinien gezogen.

Es wurden, bei einer Pufferkonzentration von 0.75 M, die pH-Werte 8.5 (Imidazol-Puffer), 8.0 (Imidazol-Puffer), 7.0 (Triethylamonium-acetat) und 5.0 (2-(*N*-Morpholino)ethansulfonsäure) verglichen. Leider wurden die Oligonukleotide bei dem pH-Wert von 5.0 zersetzt und es konnte keine Aussage über die Umsetzung gemacht werden. Wie schon beim Einsatz von *p*-cIBX waren niedrigere pH-Werte für den Umsatz nachteilig. Der maximale Umsatz wurde nach wie vor mit einem pH von 8.5 (Imidazol-Puffer) erreicht (Abbildung 63).

Zusammenfassend konnte eine erfolgreiche Synthese der positiv geladenen IBX-Derivate **39** und **27** beschrieben werden. Beide Moleküle wurden an hmdC im Kontext des Test-Oligonukleotids ODN **17** auf ihre selektive Oxidationsfähigkeit hin zum fdC getestet, mit maximal 8.8 % (**39**) bzw. 11 % (**27**) fdC. *p*-cIBX **22** ist im Vergleich mit 70-80 % Umsatz zu fdC deutlich besser. Ein Grund für die höhere Oxidationskraft von **22** könnte in der Carboxylgruppe liegen. Der leicht elektronenziehende Effekt dieser Gruppe könnte den Angriff des Alkoholsauerstoffs am, damit elektronenärmeren, Iod erleichtern. Ein weiterer Grund könnte die beobachtete Instabilität der IBX-Derivate **39** und **27** bei 60 °C sein. Diese Instabilität könnte durch das flexible PEG Gerüst verursacht werden. Es gibt Hinweise darauf, dass die Oxidation mit IBX im wässrigen Milieu radikalisch abläuft.^[229] Aufgrund dessen, könnte das Molekül, wie für Anilide bereits beschrieben,^[233, 236] intramolekular angreifen. Eine starrere Struktur als die hier verwendete Struktur könnte dies umgehen, aber würde auch die Bindung an das Phosphatrückgrat einschränken.

6. Die Modifikation ms^2A in RNA

Für den Ursprung von ms^2A in humanen Zellen gibt es prinzipiell zwei Möglichkeiten (Abbildung 64). Zum einen könnte ms^2A direkt mit Hilfe einer Thiomethyl-Transferase als eigenständige Modifikation in RNA eingebaut werden. Zum anderen könnte ms^2A über eine aktive und teilweise Demodifizierung von 2-thiomethylierten Adenosin-Modifikationen entstehen. In humanen Zellen konnten bisher nur ms^2i^6A und ms^2t^6A nachgewiesen werden.^[112]

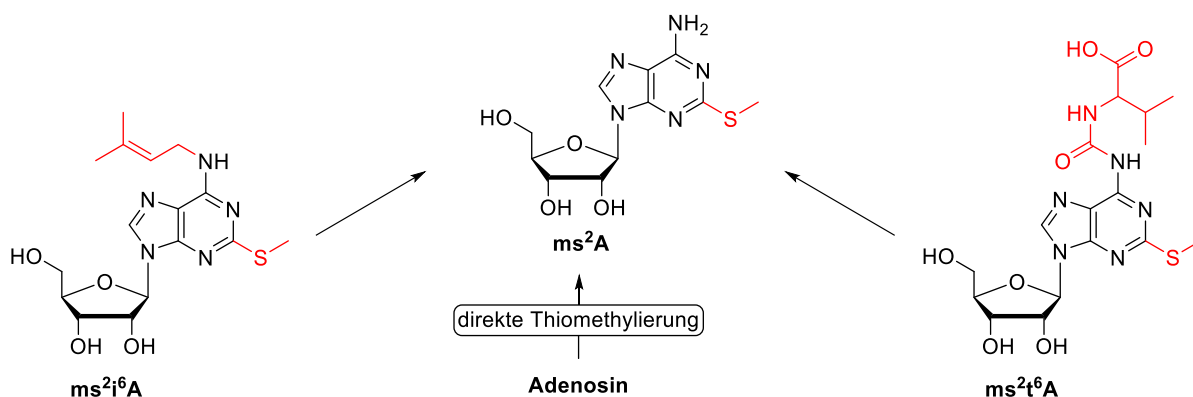


Abbildung 64: Mögliche Biosynthesewege der RNA Modifikation ms^2A .

Daher sollte ein an der Thiomethyl-Gruppe dreifach deuteriertes ms^2A ($d3-ms^2A$) synthetisiert werden, um es in quantitativen LC-MS/MS Studien einzusetzen.

6.1. Synthese eines $d3-ms^2A$

Die Synthese des $d3-ms^2A$ wurde ausgehend von Guanodin **40** in vier Schritten durchgeführt und die Markierung mit Deuterium im vorletzten Schritt eingeführt (Abbildung 65).^[237]

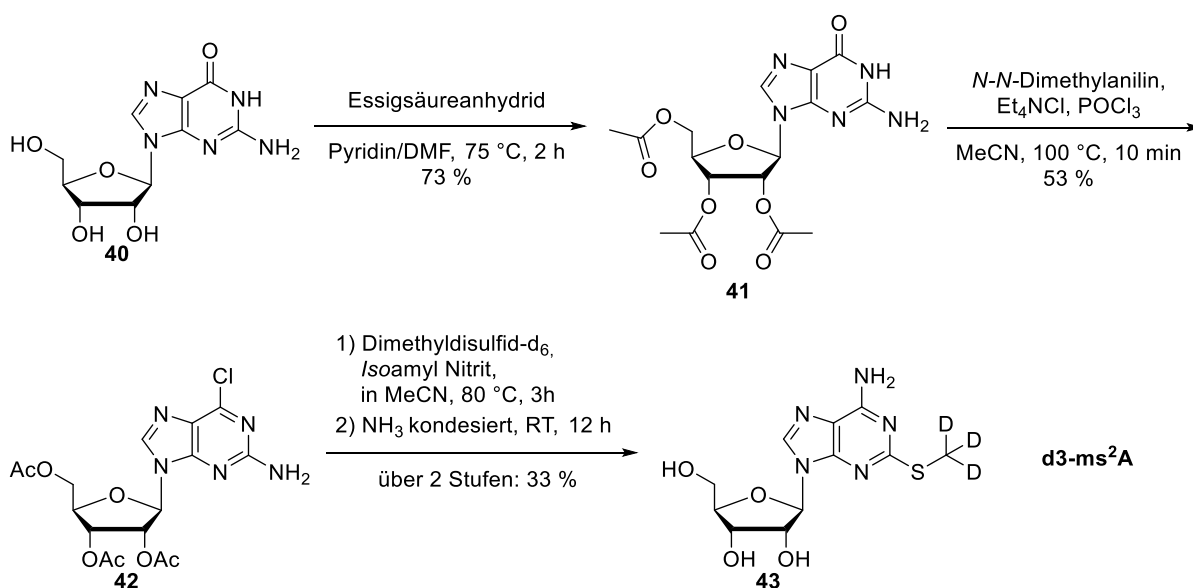


Abbildung 65: Syntheseroute zur Darstellung des $d3-ms^2A$ **43**, ausgehend von Guanodin **40**.

Das Guanosin **40** wurde zunächst mittels Essigsäureanhydrid an den freien Hydroxygruppen geschützt. Das Acetyl-geschützte Guanosin **41** wurde mit *N,N*-Dimethylanilin, $POCl_3$ und Tetramethylammoniumchlorid umgesetzt, um im letzten Schritt die Amin-Gruppe in Position 6 einführen zu können. $POCl_3$ reagiert dabei, ähnlich wie bei der Formierung des Vilsmeier Reagenzes (Vilsmeier-Haack Reaktion), mit dem Sauerstoff der Guaninbase und formt die Chlorospezies **42**. Im folgenden Schritt wurde über d6-Dimethyldisulfid die Deuteriummarkierung eingeführt. Dafür wurde mittels *Isoamyl*nitrit die 2-Amin-Gruppe von **42** in eine hoch reaktive Diazogruppe überführt, welche zur d3-Thiomethylgruppe weiterreagiert. Dabei wurde das Chloropurin **42** im Überschuss verwendet, um das eingesetzte d6-Dimethyldisulfid möglichst quantitativ umzusetzen. Das Produkt wurde isoliert und direkt mit kondensiertem Ammoniak in einem Druckrohr umgesetzt. Das Lösungsmittel Ammoniak hatte dabei zwei weitere Funktionen: Zum einen reagierte es mit dem 6-Chloropurin zum exozyklischen N^6 -Amin und zum anderen wurden durch das basische Milieu die Hydroxylgruppen des Zuckers entschützt. Nach zweifacher chromatographischer Auftrennung mittels RP-HPLC konnte **43** in reiner Form erhalten werden.

6.2. Anwendung des synthetisierten Standards

Als erster Modellorganismus wurde das Bakterium *E. coli* ausgewählt, da die Modifikation ms^2A dort, im Vergleich zu humanen Zellen, abundant ist.^[115] Zudem verfügt dieser Organismus nur über ms^2i^6A als mögliche direkte Vorläuferverbindung von ms^2A .^[162] Mit Hilfe des in dieser Arbeit synthetisierten d3- ms^2A konnte *Timm Ensfelder* in *E.coli* bestätigen, dass ein direkter Zusammenhang zwischen der Hypermodifikationen ms^2i^6A und ms^2A zu bestehen scheint (Dissertation von *Timm Ensfelder*).

Aufgrund des erfolgreichen Experiments mit *E.coli*, wurde ein vergleichbares Experiment mit menschlichen HAP1-Zellen (*human near-haploid* Zellen abgeleitet von KBM7-Zellen) durchgeführt. In Zusammenarbeit mit *Timm Ensfelder* wurden, zur Verfügung stehende, HAP1-Zellen mit einem CDK5RAP1 *Knockout* kultiviert. Das Enzym CDK5RAP1 überträgt die Thiomethyl-Gruppe auf das Substrat i^6A . Daher sollte die Gesamt-RNA der CDK5RAP1-KO kein ms^2i^6A aufweisen. Die Zellen wurden lysiert und die Gesamt-RNA isoliert und von *Mirko Wagner* quantifiziert. Übereinstimmend mit den Daten von *Kellner et al.* konnten wir ms^2A in RNA von humanen Zellen nachweisen. In Abbildung 66 ist das extrahierte Ionen-Chromatogramm (EIC) der massenspektrometrischen Messungen von ms^2A zu sehen (Bereich für ms^2A (m/z): 314.0868-314.0968; Bereich für d3- ms^2A (m/z): 317.1056-317.1156). ms^2A liegt in den HAP1-Zellen zwar nur in geringen Mengen vor, kann aber zweifelsfrei als Signal identifiziert werden.

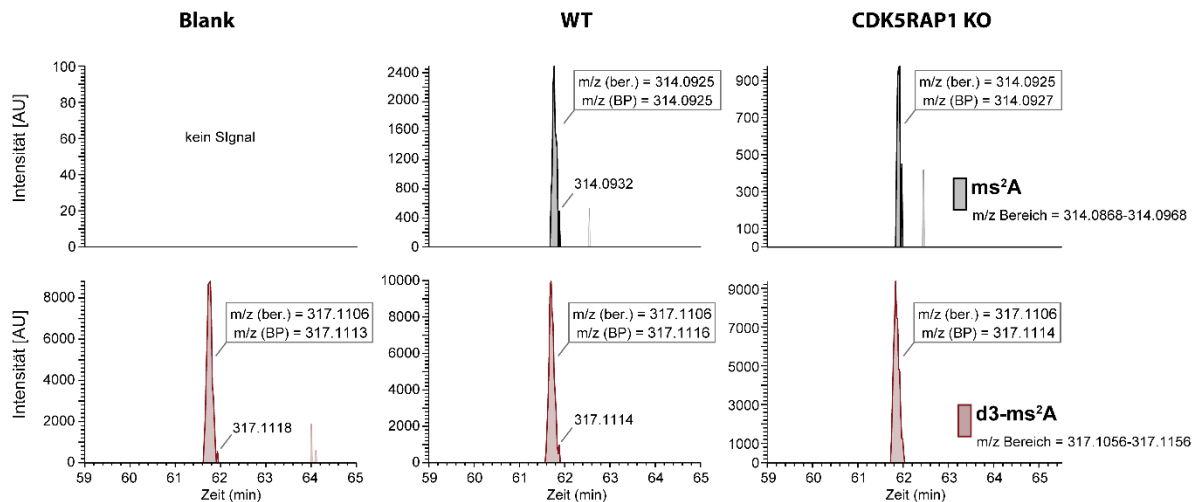


Abbildung 66: Ein extrahiertes Ionen-Chromatogramm (EIC) der massenspektrometrischen Quantifizierung von ms^2A in HAP1-WT und HAP1-CDK5RAP1-KO. Die linken Chromatogramme zeigen die Messungen der Negativkontrolle des Verdau (Blank), die mittleren Chromatogramme zeigen die Messungen des WT und die rechten Chromatogramme zeigen die Messungen des CDK5RAP1-KO. Die oberen Chromatogramme (schwarze Signale) zeigen jeweils das natürliche ms^2A . Die unteren Chromatogramme (rote Signale) zeigen jeweils den hinzugefügten (*spiked*) synthetisierten d3- ms^2A Standard.

In Abbildung 67 werden alle gemessenen Mengen an i^6A , ms^2i^6A und ms^2A aus dem HAP1-WT und des HAP1-CDK5RAP1-KO zusammengefasst dargestellt. Der CDK5RAP1-KO zeigte wie erwartet kein ms^2i^6A und eine höhere Menge an i^6A im Vergleich zum WT. Im Einklang mit der eingangs beschriebenen Hypothese, dass ms^2A ein direktes Abbauprodukt von ms^2i^6A sein könnte, konnte auch im CDK5RAP1-KO ein Rückgang um mehr als die Hälfte der Menge an ms^2A beobachtet werden. Diese Beobachtung unterstützt die Vermutung, dass ms^2A in *E. coli* zu einem großen Teil aus ms^2i^6A gebildet wird. Allerdings weisen die Daten darauf hin, dass der Weg über ms^2i^6A nicht der alleinige Entstehungsweg zu sein scheint, da ms^2A nicht vollständig verschwindet.

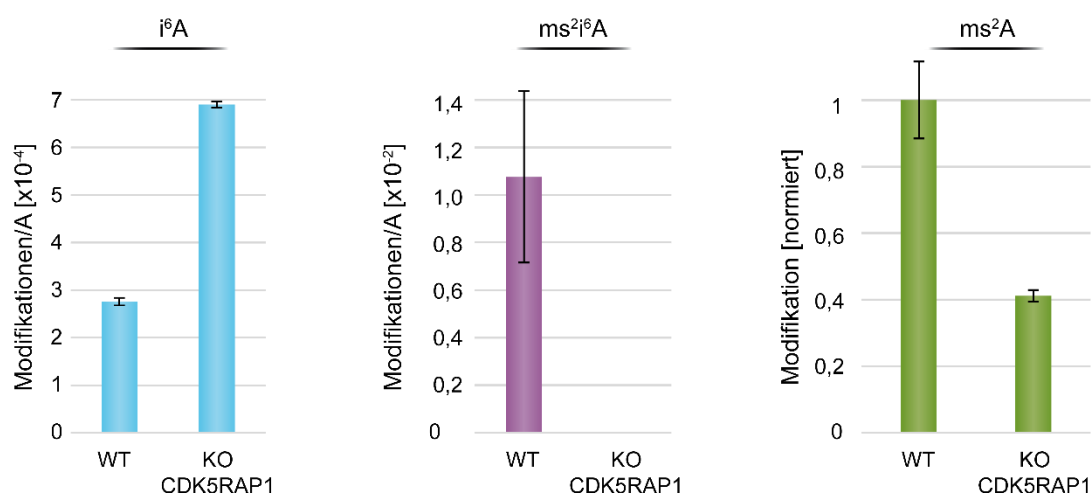


Abbildung 67: Quantifizierungsdaten der Modifikationen i^6A , ms^2i^6A und ms^2A , von isolierter Gesamt-RNA aus einem HAP1 WT und dessen CDK5RAP1 KO. Die i^6A Daten (blau) und ms^2i^6A Daten (lila) wurden im Verhältnis zu der Menge an Adenosin aufgetragen. Die Modifikation ms^2A (grün) wurde normiert aufgetragen.

Damit konnten gezeigt werden, dass ms^2i^6A mit großer Wahrscheinlichkeit auch in menschlichen Zellen ein Vorläufer von ms^2A ist. Der Abbau von ms^2i^6A könnte zum einen den normalen Vorgang während des Metabolismus der tRNA darstellen, oder ein Hinweis, auf die erste reversible Hypermodifikation, sein.

Um die biologische Wirkung von ms^2A zu untersuchen, könnten im weiteren Schritt Interaktomstudien durchgeführt werden, wie sie *Geiger et al.* erfolgreich für m^6A durchgeführt haben.^[238] Durch diese Studien könnten spezifische *reader* und *repeller* Proteine identifiziert werden. So könnte möglicherweise auch das für die Biosynthese von ms^2A zuständige Protein identifiziert werden, da es vermutlich als ein *repeller* Protein agiert. Aber auch Informationen über mögliche Funktionen der Modifikation oder der mögliche weitere Abbau zu Adenosin könnten damit gesammelt werden.

7. Veröffentlichte Arbeit

ALKBH5-induzierte Demethylierung von mono und dimethyliertem Adenosin

„ALKBH5-induced demethylation of mono and dimethylated adenosine“ Timm T. Ensfelder[#], Matthias Q. Kurz[#], Katharina Iwan[#], Simon Geiger, Sarah Matheisl, Markus Müller, Roland Beckmann und Thomas Carell, *Chem. Commun.*, **2018**, 54, 8591-8593, DOI: 10.1039/c8cc03980a.

Die mit [#] markierten Autoren haben zu gleichen Teilen beigetragen.

Prolog

Ähnlich der epigenetischen Informationsebene, eröffnet die katalysierte und aktive Demethylierung von m⁶A durch ALKBH5 die Möglichkeit einer transkriptomischen Informationsebene. Wir beschäftigten uns in der vorliegenden Studie mit der Frage, ob ALKBH5 auch das in rRNA natürlich vorkommende m⁶₂A demethylieren kann, trotz des dimethylierten exozyclischen N⁶-Heteroatoms. Die Modifikation m⁶₂A befindet sich ausschließlich in der humanen 40S Untereinheit des Ribosoms (18S rRNA; Position 1850 und 1851) in einem hoch konservierten Sequenzkontext.^[121] Die Dimethylierung der beiden aufeinander folgenden Positionen 1850 und 1851 zu m⁶₂A ist dabei essentiell für ein korrektes *ribosome assembly*.^[239] Eine Umsetzung von m⁶₂A mit ALKBH5 könnte zusätzlich tiefere Einblicke in den Mechanismus der Demethylierung bringen. Möglicherweise können, mit der vermuteten Verlangsamung der Demethylierung durch m⁶₂A, Zwischenprodukte auftreten, die das m⁶A nicht zeigt.

In ersten, von mir durchgeführten Aktivitäts-Experimenten, konnten tatsächlich Hinweise auf eine Aktivität von ALKBH5 gegenüber m⁶₂A im Sequenzmotiv von m⁶A (GGACU) gezeigt werden. Dafür wurden Oligonukleotide mit der Sequenz 5'-GGm⁶₂ACU-3' synthetisiert und mit ALKBH5, α-Ketoglutarat und einem Fe²⁺-haltigem Puffer umgesetzt.^[159]

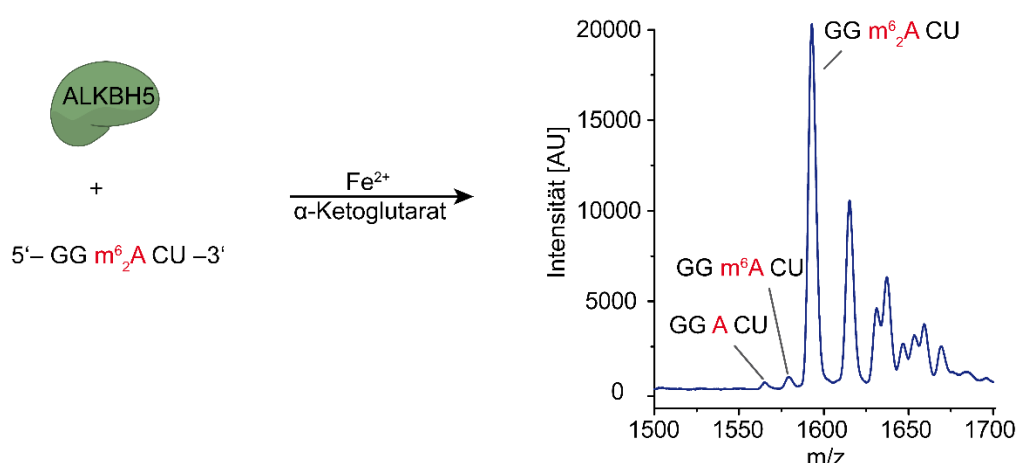


Abbildung 68: Darstellung der ersten Aktivitäts-Experimente von ALKBH5 mit der Modifikation m⁶₂A im Sequenzmotiv von ALKBH5 (5' - GGm⁶₂ACU - 3'). Auf der rechten Seite ist ein Massenspektrum (MALDI-TOF) gezeigt (blau). Die Signale für das eingesetzte Oligonukleotid, die einfach demethylierten Spezies (-14 Da) und die zweifach demethylierte Spezies (-28 Da) wurden markiert.

Interessanterweise konnten im erzeugten MALDI-Massenspektrum (Abbildung 68) eine einfache Demethylierung zu 5'-GGm⁶ACU-3' (-14 Da im Vergleich zum eingesetzten Oligonukleotid) und eine zweifache Demethylierung zu 5'-GGACU-3' (-28 Da) beobachtet werden. Beflügelt durch diese positiven Ergebnisse gingen wir noch einen Schritt weiter und synthetisierten das natürliche Sequenzmotiv von m⁶₂A (5'-GUGm⁶₂Am⁶₂ACUU-3'), um auch hier die Demethylierung durch ALKBH5 zu untersuchen. Als Vergleich dienten Oligonukleotide mit m⁶A in den jeweiligen Sequenzmotiven. *Timm Ensfelder* übernahm die Optimierung der ALKBH5 Expression und der Reaktionsbedingungen.

Wir konnten zeigen, dass ALKBH5 die Modifikation m⁶₂A, sowohl in rRNA Kontext, als auch in kleinen Oligonukleotiden, demethylieren kann. Im konservierten Sequenzkontext von m⁶₂A konnten unterschiedliche Intermediate von der mono-demethylierten bis zur vollständig demethylierten Spezies beobachtet werden. Zudem konnten wir zeigen, dass ALKBH5 m⁶A in einem weitaus größeren Sequenzspektrum demethylieren kann, als bisher angenommen, wenn es equimolar oder in höheren Verhältnissen eingesetzt wird.

Autorenbeitrag

In dieser Arbeit war ich für die Durchführung der Synthesen der m⁶A Phosphoramidite und der zu untersuchenden m⁶A Oligonukleotide zuständig. Die Oligonukleotide wurden mittels RP-HPLC von mir gereinigt. Die Orbitrap-MS Experimente und deren Auswertung führte ich mit Unterstützung von *Christina Schneider* durch. Alle Experimente und erhaltenen Daten wurden in enger Zusammenarbeit mit *Timm Ensfelder*, *Katharina Iwan*, *Simon Geiger*, *Sarah Matheisl*, *Markus Müller*, *Roland Beckmann* und *Thomas Carell* geplant und diskutiert.

Lizenz

Kopie der Publikation mit Erlaubnis des Verlags Royal Society of Chemistry.



Cite this: *Chem. Commun.*, 2018, 54, 8591

Received 18th May 2018,
Accepted 4th July 2018

DOI: 10.1039/c8cc03980a

rsc.li/chemcomm

ALKBH5-induced demethylation of mono- and dimethylated adenosine†

Timm T. Ensfelder,^{‡a} Matthias Q. Kurz,^{§a} Katharina Iwan,^{§a} Simon Geiger,^a Sarah Matheisl,^b Markus Müller,^{§a} Roland Beckmann^b and Thomas Carell^{§*a}

RNA contains methylated A-base derivatives. A methylation to m⁶A and then demethylation regulate homeostasis in mRNA. It is assumed that m⁶A is mainly demethylated by the α -ketoglutarate dependent oxidase ALKBH5. Here we show that ALKBH5 also demethylates the dimethylated adenosine m⁶₂A, which is a non-canonical base present in ribosomal RNA.

RNA contains a large number of modified nucleosides.^{1,2} Most abundant throughout all RNA species are mono-methylated or multiply methylated nucleosides.² Particularly high levels of methylated nucleosides are detected in transfer RNA (tRNA), but also ribosomal RNA (rRNA) and messenger RNA (mRNA) feature such non-canonical nucleosides.^{1,3} Basically, for all canonical nucleosides methylated derivatives exist.¹ They all serve different functions.⁴ The methylated RNA nucleoside m⁶A was first reported in mRNA in 1974^{5,6} and is among the best studied modifications within the set of methylated adenosines (Fig. 1).⁷ Others^{8–14} and we¹⁵ have shown that the nucleoside recruits and repels proteins to control the lifetime of the corresponding mRNA. As such the nucleoside is directly involved in regulating mRNA homeostasis.¹⁵ Particularly exciting is the observation that the nucleoside m⁶A can be demethylated by α -ketoglutarate dependent oxidases,¹⁶ which opens the possibility that methylation and demethylation establish a new layer of regulation to facilitate transcriptional control. In addition to m⁶A, ribosomal RNA also contains the double methylated species m⁶₂A, where two methyl groups are both situated at the exocyclic N⁶-heteroatom. Because the methyl groups are connected to a heteroatom, this oxidation would allow oxidative double demethylation of m⁶₂A in contrast to m²A and m⁸A, where the methyl

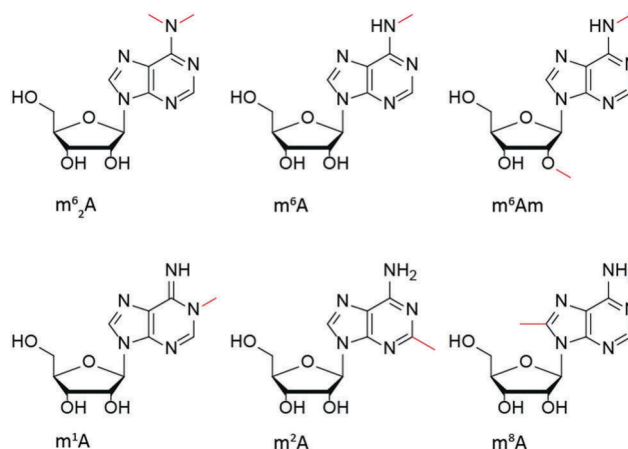


Fig. 1 Selection of naturally occurring methylated adenosines in RNA.

group resides on a carbon. Such sequential demethylations performed by a single enzyme are known: for example, histone demethylases are able to demethylate mono-, di- and even trimethylated lysines.^{17,18} Here, we investigated if the Fe²⁺-, α -ketoglutarate-dependent ALKBH5, which is known to demethylate m⁶A, would be able to demethylate naturally occurring m⁶₂A as well. In contrast to m⁶A, however, m⁶₂A is predominately, if not exclusively, found in a conserved sequence context of the small ribosomal subunit.^{2,3} For this study, we first investigated the demethylation of small RNA strands containing m⁶A and m⁶₂A in different sequence contexts with ALKBH5_{66–292} *in vitro*. Therefore, we synthesized the phosphoramidites of m⁶A and m⁶₂A and subsequently used them for solid phase RNA synthesis (SPOS) of oligoribonucleotides (ORN) 1–4 (Fig. 2). In this study we inserted m⁶A and m⁶₂A into their natural sequence context, *i.e.*, 5'-GGm⁶ACU-3' for m⁶A and 5'-GUGm⁶₂A-m⁶ACUU-3' for m⁶₂A. To address potential sequence bias during the enzymatic studies, we also inserted m⁶A into the sequence of m⁶₂A and *vice versa*. In addition to the investigation with pure oligonucleotides, the study was furthermore complemented with the putative demethylation of m⁶₂A in intact ribosomes and in human ribosomal rRNA.

^a Center for Integrated Protein Science, Department of Chemistry, Ludwig-Maximilians-Universität München, Butenandtstraße 5-13, 81377 Munich, Germany. E-mail: Thomas.carell@lmu.de; Web: www.carellgroup.de

^b Gene Center and Department of Biochemistry, Ludwig-Maximilians-Universität München, Feodor-Lynen-Straße 25, 81377 Munich, Germany

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8cc03980a

‡ Equal contributions.

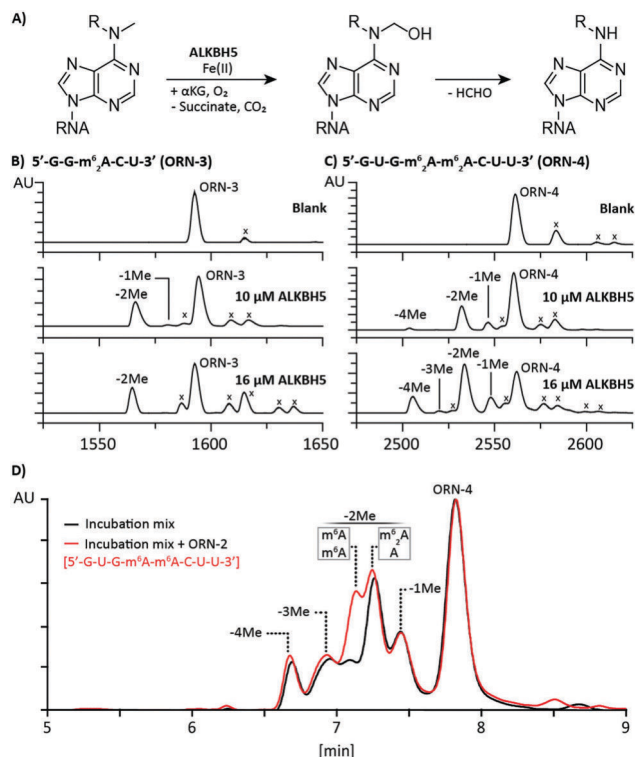


Fig. 2 (A) Oxidation scheme of methylated adenine by ALKBH5 with R being a hydrogen or a methyl group. (B and C) MALDI-TOF spectra of ORN 3–4 and their demethylation products ($-x\text{Me}$) at different concentrations of ALKBH5_{66–292}. The sodium signals in the spectra are marked with an “x”. (D) HPL chromatogram of the incubation of ORN-4 with 16 μM ALKBH5_{66–292} (black line) and with a spike of ORN-2 (red line).

The His₆-tagged protein human ALKBH5_{66–292} enzyme was overexpressed in *E. coli* BL21(DE3) from a pNIC28-Bsa4-ALKBH5_{66–292} plasmid¹⁹ using a modified procedure compared to the literature²⁰ data (see the ESI†) and isolated in one step *via* Ni-NTA affinity chromatography. This procedure yielded the protein in a purity of 90% (ESI,† Fig. S1). The buffer was changed to the assay buffer (50 mM Tris HCl, pH 7.5, 15 mM KCl). Stock solutions of diammonium-Fe(II)-sulfate and α -ketoglutarate were prepared fresh for every assay and added to a final concentration of 300 μM . ALKBH5_{66–292} concentrations in the assay buffer were either 10 or 16 μM . The corresponding oligonucleotides were held constant at 10 μM . Exclusion of oxygen was not needed. The solution was incubated at 37 °C for about 1 h (see the ESI†). For analysing the demethylation reaction, we desalted the assay solution using a Millipore MF™ VSWP-membrane (0.025 μm) and measured the molecular weight of the oligonucleotides using MALDI-TOF mass spectrometry. The obtained data are depicted in Fig. 2.

As expected, we observed complete demethylation of m⁶A in both sequence contexts ORN-1 and ORN-2 (ESI,† Fig. S2A and B). This is interesting because it shows that the oxidation of m⁶A, putatively to the corresponding N⁶-hydroxymethyl-A followed by hydrolysis of the hemiaminal (Fig. 2A), has no observable sequence bias under our conditions. In addition, we see that the demethylation proceeds also at two directly adjacent m⁶A to completion.

When we analysed the demethylation with the m⁶₂A strands ORN-3 and ORN-4 (Fig. 2B and C), we obtained more complex spectra. We could indeed detect demethylation but with a significantly reduced efficiency. When m⁶₂A is placed in the context of the m⁶A consensus sequence, the formation of the mono-demethylated ORN-3 at 10 μM enzyme concentration is seen, furthermore the fully demethylated ORN-3 (-2Me) is clearly visible (Fig. 2B). The conversion is about 40%, estimated based on the observed signal intensities. Demethylation in the rRNA context, where two directly adjacent m⁶₂A occur, is shown in Fig. 2C. At 10 μM and an incubation time of 1 h, mono-, di-, and tetra-demethylations are observed. At a slightly higher enzyme concentration, we detected all expected species with the loss of one–four methyl groups. The loss of two methyl groups can yield two different species, either one with m⁶A–m⁶A or one with m⁶₂A–A. To address this issue, we further performed LC-MS analysis of the ORN-4 assay mixture with and without a spike of ORN-2. Fig. 2D shows that both species occur and the m⁶₂A–A product is the favoured one. In summary, the data show that ALKBH5_{66–292} is able to fully demethylate an m⁶₂A–m⁶₂A sequence *via* oxidation of all four methyl groups. Interestingly, the mono- and tri-demethylated ORNs do not accumulate. We interpret this effect as an indication that the first demethylation step is rate limiting, while the demethylation of m⁶A proceeds quickly.

We next wanted to study if double-demethylation of m⁶₂A is also possible on intact rRNA. To this end, we isolated the ribosomes from HEK293T cells using ultracentrifugation followed by rRNA isolation. The obtained ribosomes and rRNA are more complex than the synthesized oligonucleotides because they contain one m⁶A per subunit and in addition the m⁶₂A–m⁶₂A sequence. These m⁶₂A positions are highly conserved in all species with only very few exceptions.²¹ For the analysis, due to the large size of the RNA strand, we digested the RNA after the assay to the individual nucleosides and measured all four canonical nucleosides plus the methylated adenosines m⁶A and m⁶₂A *via* UHPLC-MS/MS using our described procedure (see the ESI†). In order to obtain absolute quantitative data, we synthesized the isotopologues of m⁶A and m⁶₂A that were used for internal calibration (see the ESI†). The obtained quantitative MS-data are compiled in Fig. 3. In the control studies performed without enzymes, we clearly detected the m⁶A and m⁶₂A nucleosides at the expected ratios (two of each per rRNA). Upon addition of ALKBH5_{66–292} in two different concentrations, we see a clear and strong reduction of the m⁶A signal, showing that in principle ALKBH5_{66–292} is able to demethylate m⁶A also in ribosomal RNA. Because the sequence context is different, this result argues again that the activity of ALKBH5_{66–292} has no observable sequence preference. One has to note that in these experiments ALKBH5_{66–292} is present in great excess over the used large rRNA. Interestingly, however, m⁶A levels are not reduced to zero which we would expect based on the observed full demethylation with the smaller RNA strands. For m⁶₂A, we also see clearly reduced levels, showing that ALKBH5_{66–292} demethylated m⁶₂A also in rRNA. It is interesting to note that at 16 μM enzyme concentration, in contrast to the ORN, the resulting m⁶A values are higher than those observed at

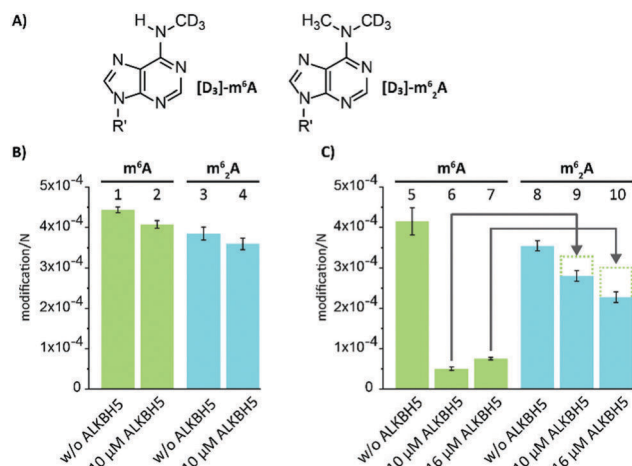


Fig. 3 Quantification data of m⁶A and m⁶₂A upon incubation with ALKBH5_{66–292} at different concentrations. (A) For quantification, isotope labelled standards of m⁶A and m⁶₂A were used. R' represents the ribose sugar moiety. Panel (B) represents the amount in ribosomes and (C) the amount in rRNA. The dashed lines shown in (C) indicate the sum of residual m⁶A and m⁶₂A.

10 μM, despite lower m⁶₂A levels. We hypothesize that the observed m⁶A levels are the hemi-demethylated intermediates that are obtained when ALKBH5_{66–292} oxidizes and demethylates m⁶₂A. The sum of residual m⁶A and m⁶₂A levels never exceeds the starting values of m⁶₂A (Fig. 3D), showing that the full demethylation of m⁶₂A occurs and that all the original m⁶A was likely demethylated as well.

ALKBH5 is an α-ketoglutarate-dependent oxidase that was shown to demethylate m⁶A in mRNA. We investigated here if the enzyme is able to demethylate also the double-methylated m⁶₂A, which is known to occur in the small subunit of the ribosome. We also wanted to know if the enzyme is in addition able to demethylate m⁶A in rRNA. We show that ALKBH5 demethylates m⁶A in basically all sequence contexts, when used at an equimolar ratio or higher, even when two m⁶A bases are adjacent to each other. It also demethylates m⁶₂A both in rRNA and in small ORNs. Here, in general demethylation occurs and the mono-demethylated intermediate (ORN3/ORN4 -1Me; Fig. 2D and E) is detectable. This is very interesting because it shows that the enzyme releases its substrate after the first demethylation step. Since the demethylation of the mono-methylated compound m⁶A is much more efficient compared to m⁶₂A, one would expect direct conversion of m⁶₂A into adenosine if the substrate is not released. We speculate that the substrate release after the first demethylation could be required to reload another α-ketoglutarate co-substrate into the active site. It is unclear if the demethylation of m⁶₂A ALKBH5_{66–292} is a process that occurs in nature. We suspect that the small amount of observed demethylation at approximately 100-fold excess of the enzyme relative to the added rRNA speaks against this possibility. Indeed, when we performed the demethylation studies with assembled ribosomes, demethylation did not occur. This makes demethylation of intact ribosomes unlikely, but it still leaves the possibility that demethylation occurs during rRNA maturation

similar to the insertion of non-canonical nucleosides into mRNA upon mRNA maturation.

We thank Felix Hagelskamp and the Kellner Lab (LMU Munich) for sharing unpublished protocols. We thank Christina Schneider for applying this protocol to help us with the LC-MS measurements. We thank Michael McDonough, Christopher Schofield and Dong Zhang for kindly providing us with the pNIC28-Bsa4-ALKBH5_{66–292} plasmid. Furthermore, we thank the Deutsche Forschungsgemeinschaft (SFB749, SFB1032, GRK2062 and SPP1784) for financial support. Additional support was obtained from the Excellence Cluster CIPSM.

Conflicts of interest

There are no conflicts to declare.

Notes and references

- 1 P. Boccaletto, M. A. Machnicka, E. Purta, P. Piątkowski, B. Bagiński, T. K. Wirecki, V. de Crécy-Lagard, R. Ross, P. A. Limbach, A. Kotter, M. Helm and J. M. Bujnicki, *Nucleic Acids Res.*, 2018, **46**, D303–D307.
- 2 M. A. Machnicka, K. Milanowska, O. Osman Oglou, E. Purta, M. Kurkowska, A. Olchowik, W. Januszewski, S. Kalinowski, S. Dunin-Horkawicz, K. M. Rother, M. Helm, J. M. Bujnicki and H. Grosjean, *Nucleic Acids Res.*, 2013, **41**, D262–D267.
- 3 D. Piekna-Przybylska, W. A. Decatur and M. J. Fournier, *Nucleic Acids Res.*, 2008, **36**, D178–D183.
- 4 G. Jia, Y. Fu and C. He, *Trends Genet.*, 2013, **29**, 108–115.
- 5 R. Desrosiers, K. Friderici and F. Rottman, *Proc. Natl. Acad. Sci. U. S. A.*, 1974, **71**, 3971–3975.
- 6 R. P. Perry and D. E. Kelley, *Cell*, 1974, **1**, 37–42.
- 7 B. S. Zhao, I. A. Roundtree and C. He, *Nat. Rev. Mol. Cell Biol.*, 2016, **18**, 31–42.
- 8 W. Xiao, S. Adhikari, U. Dahal, Y.-S. Chen, Y.-J. Hao, B.-F. Sun, H.-Y. Sun, A. Li, X.-L. Ping, W.-Y. Lai, X. Wang, H.-L. Ma, C.-M. Huang, Y. Yang, N. Huang, G.-B. Jiang, H.-L. Wang, Q. Zhou, X.-J. Wang, Y.-L. Zhao and Y.-G. Yang, *Mol. Cell*, 2016, **61**, 507–519.
- 9 D. P. Patil, C.-K. Chen, B. F. Pickering, A. Chow, C. Jackson, M. Guttman and S. R. Jaffrey, *Nature*, 2016, **537**, 369–373.
- 10 X. Wang, B. S. Zhao, I. A. Roundtree, Z. Lu, D. Han, H. Ma, X. Weng, K. Chen, H. Shi and C. He, *Cell*, 2015, **161**, 1388–1399.
- 11 K. D. Meyer, D. P. Patil, J. Zhou, A. Zinoviev, M. A. Skabkin, O. Elemento, T. V. Pestova, S.-B. Qian and S. R. Jaffrey, *Cell*, 2015, **163**, 999–1010.
- 12 X. Wang, Z. Lu, A. Gomez, G. C. Hon, Y. Yue, D. Han, Y. Fu, M. Parisien, Q. Dai, G. Jia, B. Ren, T. Pan and C. He, *Nature*, 2014, **505**, 117–120.
- 13 K. D. Meyer, Y. Saletore, P. Zumbo, O. Elemento, C. E. Mason and S. R. Jaffrey, *Cell*, 2012, **149**, 1635–1646.
- 14 D. Dominissini, S. Moshitch-Moshkovitz, S. Schwartz, M. Salmon-Divon, L. Ungar, S. Osenberg, K. Cesarkas, J. Jacob-Hirsch, N. Amariglio, M. Kupiec, R. Sorek and G. Rechavi, *Nature*, 2012, **485**, 201–206.
- 15 R. R. Edupuganti, S. Geiger, R. G. H. Lindeboom, H. Shi, P. J. Hsu, Z. Lu, S.-Y. Wang, M. P. A. Baltissen, P. W. T. C. Jansen, M. Rossa, M. Müller, H. G. Stunnenberg, C. He, T. Carell and M. Vermeulen, *Nat. Struct. Mol. Biol.*, 2017, **24**, 870–878.
- 16 G. Zheng, J. A. Dahl, Y. Niu, P. Fedorcsak, C.-M. Huang, C. J. Li, C. B. Vågbo, Y. Shi, W.-L. Wang, S.-H. Song, Z. Lu, R. P. G. Bosmans, Q. Dai, Y.-J. Hao, X. Yang, W.-M. Zhao, W.-M. Tong, X.-J. Wang, F. Bogdan, K. Furu, Y. Fu, G. Jia, X. Zhao, J. Liu, H. E. Krokan, A. Klungland, Y.-G. Yang and C. He, *Mol. Cell*, 2013, **49**, 18–29.
- 17 K. Yamane, C. Toumazou, Y.-I. Tsukada, H. Erdjument-Bromage, P. Tempst, J. Wong and Y. Zhang, *Cell*, 2006, **125**, 483–495.
- 18 S. M. Kooistra and K. Helin, *Nat. Rev. Mol. Cell Biol.*, 2012, **13**, 297–311.
- 19 W. Aik, J. S. Scotti, H. Choi, L. Gong, M. Demetriades, C. J. Schofield and M. A. McDonough, *Nucleic Acids Res.*, 2014, **42**, 4741–4754.
- 20 F. W. Studier, *Protein Expression Purif.*, 2005, **41**, 207–234.
- 21 Z. Xu, H. C. O'Farrell, J. P. Rife and G. M. Culver, *Nat. Struct. Mol. Biol.*, 2008, **15**, 534–536.

Supporting Information

**ALKBH5-Induced Demethylation of Mono- and Dimethylated
Adenosine**

Timm T. Ensfelder, ‡^a Matthias Q. Kurz, ‡^a Katharina Iwan, ‡^a Simon Geiger,^a Sarah Matheisl,^b
Markus Müller,^a Roland Beckmann,^b Thomas Carell,^{*a}

Center for Integrated Protein Science (CiPS^M) at the Department of Chemistry, Ludwig-
Maximilians-Universität München, Butenandtstrasse 5–13, 81377 Munich.

Gene Center and Department of Biochemistry, Ludwig-Maximilians-Universität München,
Feodor-Lynen-Straße 25, 81377 Munich.

E-mail: Thomas.Carell@cup.uni-muenchen.de

Index

1. General methods.....	3
2. RNA-oligonucleotide solid Phase Synthesis	3
RNA oligonucleotide sequences.....	4
3. Biochemical methods	4
Ribosome isolation by Ultracentrifugation and rRNA isolation	4
Protein expression and purification	4
ALKBH5 enzyme assay	5
4. Analytical methods	6
RNA digestion	6
UHPLC/MS-MS analysis of RNA samples.....	6
Determination of extinction coefficients	7
Quantification of nucleosides.....	7
5. Supplementary Tables.....	9
6. Literature	11

1. General methods

Acetonitrile for HPLC-purification of nucleoside standards and oligonucleotides were purchased from VWR. Triethylamine and acetic acid (glacial) for HPLC buffer preparation was purchased from *Sigma-Aldrich* resp. from *Fisher Scientific*. Water of LC-MS grade was purchased from *Honeywell*. Acetonitrile of LC-MS grade was purchased from *Carl Roth GmbH + Co. KG*. Formic acid was purchased from *Fluka*, p.a. for mass spectrometry. Water was purified by an arium® pro DI Ultrapure Water System from *Satorius Stedim biotech*. Nuclease S1 (*Aspergillus oryzae*) was obtained from *Sigma Aldrich*, snake venom phosphodiesterase I (*Crotalus adamanteus*) from *USB corporation* and antarctic phosphatase from *New England Biolabs*.

2. RNA-oligonucleotide solid Phase Synthesis

The phosphoramidites of m⁶A and m⁶₂A were synthesised according to literature^{1, 2}

The m⁶A and m⁶₂A containing RNA oligonucleotides were synthesised on an ABI 394 DNA/RNA Synthesizer (*Applied Biosystems*) using typical reagent concentrations [deprotection: 3% dichloroacetic acid in DCM (*emp biotech*), activator: 0.3 M benzylthiotetrazole in MeCN (*Link Technologies*), capping: Capping A (ABI) & Capping B (ABI) solution (*J.T. Baker*), oxidation: Oxidizing (ABI) solution (*J.T. Baker*)]. The oligonucleotide synthesis were carried out on 1 µmol scale with CPG carriers (*SynBase* CPG 1000/110) and 0.1 M solutions in MeCN of TBDMS-protected standard RNA phosphoramidites (U, dmf-G, Ac-C and Bz-A) obtained from *Link Technologies*. Both modified phosphoramidites, m⁶A and m⁶₂A were incorporated into RNA using the standard protocol but the coupling times were increased to 20 min.

After synthesis, the CPG material from the cartridges were treated with 1 mL of AMA solution (28% NH₄OH in H₂O/ 40% MeNH₂ in H₂O, 1:1) first at room temperature for 5 min and secondly at 65 °C for 5 min to ensure complete cleavage of the oligonucleotide from the resin and basic deprotection. The AMA solution was removed *in vacuo*. The residue was dissolved in DMSO (100 µL) and Et₃N·3HF was added and the mixture was heated to 65 °C for 90 min. After the addition of a NaOAc solution (3 M, pH=5.5, 25 µL) the oligonucleotides were precipitated by addition of *n*-BuOH (1 mL) and cooling to -80 °C for 12 h. The suspension was centrifuged at 4 °C with 21130 x g for 60 min to pelletise the crude oligonucleotides. Further purification and analysis of the final product was performed on a Waters HPLC system (preparative HPLC: 1525 with 2487 UV detector, analytical HPLC: Waters Alliance 2695 with PDA 2996) with Nucleodur C-18 Columns from *Macherey Nagel*. The used gradient was 0.1 M Triethylammonium acetate in water and 80% MeCN. The purified oligonucleotides were

desalted using Sep-Pak C18 Classic Cartridges (*Waters*) before further usage and the identity was confirmed by MALDI-ToF-MS.

RNA oligonucleotide sequences

ORN-1: 5'-G-G-m⁶A-C-U-3'

ORN-2: 5'-G-U-G- m⁶A - m⁶A -C-U-U-3'

ORN-3: 5'-G-G-m⁶₂A-C-U-3'

ORN-4: 5'-G-U-G- m⁶₂A - m⁶₂A -C-U-U-3'

3. Biochemical methods

Ribosome isolation by Ultracentrifugation and rRNA isolation

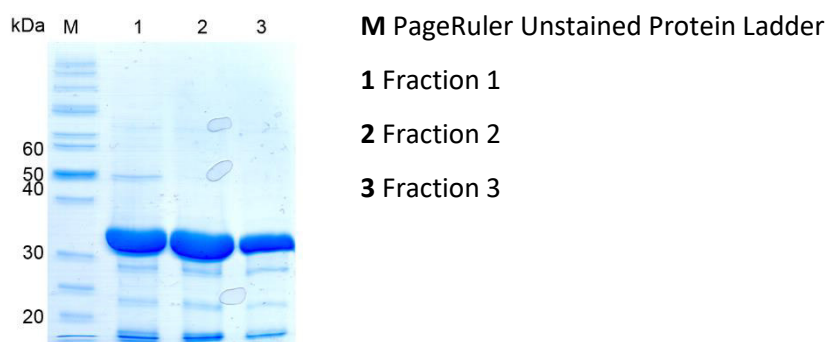
Native ribosomes were isolated from pelleted HEK 293T cells. 10·10⁷ cells were resuspended in 6 mL of lysis buffer (50 mM Tris·HCl pH 7.5, 300 mM NaCl, 6 mM MgCl₂, 0.5% NP-40 Substitute, 0.2 U/μL RNasin). Lysis was performed by incubating the cells in lysis buffer for 30 min at 4 °C. Cell debris was removed by centrifugation (*Eppendorf* FA-45-24-11 rotor, 5 min, 12000 x g, 4 °C). The supernatant was then incubated with a final Puromycin concentration of 1 mM for 15 min at RT and another 15 min at 4 °C. After the incubation the supernatant was loaded on sucrose cushion buffer (50 mM Tris·HCl pH 7.5, 300 mM NaCl, 6 mM MgCl₂, 30% (w/v) sucrose, 0.2 U/μL RNasin). The ribosomes were pelleted by centrifuging for 15 h 30 min (*Beckman Coulter* TLA 120.2, 24001 x g, 4 °C).

The supernatant was removed. The pellet was resuspended in resuspension buffer (50 mM Tris·HCl pH 7.5, 150 mM NaCl, 6 mM MgCl₂, 3.48% (w/v) sucrose) over a time of 2 h on ice. Any possible impurity was removed by another centrifugation (*Eppendorf* FA-45-24-11 rotor, 1 min, 12000 x g, 4 °C). The ribosome concentration was determined by measuring the A₂₆₀. The rRNA was isolated from the ribosome solution using the Monarch[®] Total RNA Miniprep Kit by *NEB* according to its protocol.

Protein expression and purification

A pNIC28-Bsa4 vector which is encoding an N-terminal His₆-tagged ALKBH5₆₆₋₂₉₂, was transformed into *Escherichia coli* BL21 (DE3). Expression was performed in ZYM-5052 auto induction medium without trace metals, for 3 h at 37 °C and 17 h to 20 h at 20 °C and 300 rpm. The cells were harvested by centrifugation (*Thermo Fisher Scientific* Fiberlite™ F10-4 x 1000 LEX rotor, 11970 x g, 4 °C, 8 min) and stored at -20 °C. A ratio of 10:1 (1 mL buffer per 0.1 g pellet) was used to resuspend the cell pellet in lysis buffer (50 mM Tris·HCl pH 7.5, 150 mM KCl, 10 mM Imidazol). Furthermore cOmplete™ EDTA free protease inhibitor (*Sigma-Aldrich*) and Lysozyme were added. After an incubation of 30 min on ice, the cells were lysed by sonification on ice. The cell debris was removed by centrifugation (*Thermo Fisher Scientific* Sorvall SS-34 rotor, 38720 x g, 4 °C, 30 min). The lysate was filtered and applied to a HisTrap HP (5 mL) by *GE Healthcare*. To remove unspecific binding proteins the column was washed with 8%

elution buffer (50 mM Tris-HCl pH 7.5, 150 mM KCl, 365 mM imidazole). Elution was performed with a step gradient over ten column volumes. Purification was evaluated with a 15% SDS-PAGE (Marker *Thermo Scientific* PageRuler Unstained Protein Ladder #26614). On average the purity was at 90% after one step. The buffer was changed to the assay buffer (50 mM Tris-HCl pH 7.5, 15 mM KCl) via dialysis overnight. After dialysis the enzyme solution was centrifuged for 5 min at 21130 x g and 4 °C to remove precipitated protein. The enzyme concentration was evaluated with Bradford.

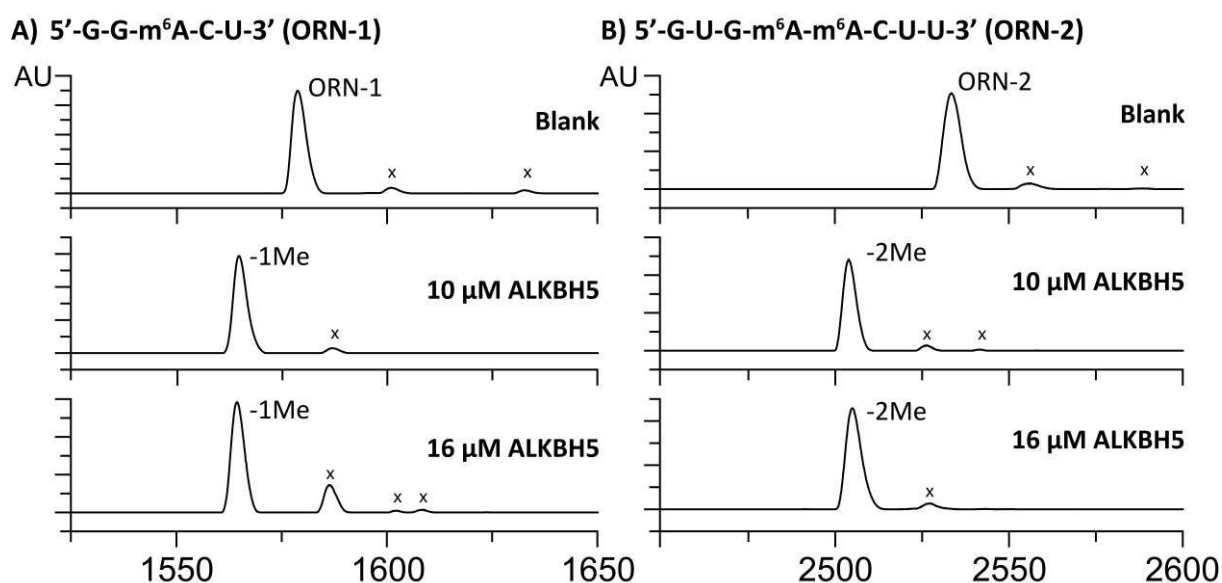


Supplementary Figure 1 ALKBH5 after one purification step. Fraction 2 was used for the enzyme assay with ORN-1 to ORN-4. Estimated purity is 90%.

ALKBH5 enzyme assay

ALKBH5₆₆₋₂₉₂ was always used directly without freezing it for storage. Every assay was performed in a volume of 50 µL and as a triplicate. Only the assay with rRNA and 16 µM ALKBH5 were performed as a duplicate. The final assay mixture was 50 mM Tris-HCl pH 7.5, 15 mM KCl, 2 mM L-ascorbate, 300 µM α-ketoglutarate, 300 µM diammonium iron(II) sulfate hexahydrate complex. Assays with native ribosomes did contain 150 mM KCl instead of 15 mM and also contain sucrose and MgCl₂ with a final concentration of 3.48% (w/v) and 6 mM. In all assays with rRNA, a murine RNase inhibitor was added with a final concentration of 1 U/µL. The L-ascorbate, α-ketoglutarate and diammonium iron(II) sulfate hexahydrate solutions were made afresh. The concentration of the synthetic RNA oligonucleotides (ORN 1-4) was 10 µM. The rRNA concentration was 101.5 nM (11.9 µg) and ribosome concentration was 250 nM. The enzyme concentration was either 10 µM or 16 µM ALKBH5₆₆₋₂₉₂.

All assays with ORNs or rRNA were incubated at 37 °C for 1 h. Assays with native ribosomes were incubated in a Thermomixer for 4 h at 37 °C and 1000 rpm. Assays with rRNA and native ribosomes were immediately stored at -20 °C for further RNA extraction. Every assay with synthetic RNA oligonucleotides was directly desalted (up to 30 min) on a *Merck Millipore* MF™ VSWP-membrane (0.025 µM) for an analysis by matrix-assisted laser desorption ionization (MALDI) mass spectrometry. A 3-Hydroxypicolinic acid (3-HPA) matrix was used. The native ribosomes and rRNA extraction from the assays was done with the Monarch® Total RNA Miniprep Kit by *NEB*.



Supplementary Figure 2 (A and B) MALDI-ToF spectra of ORN-1 and ORN-2 and their demethylation products (-xMe) at different concentrations of ALKBH5₆₆₋₂₉₂. The sodium signals in the spectra are marked with an "x".

4. Analytical methods

RNA digestion

1.0 μg of RNA in 35 μL H₂O were digested as follows: An aqueous solution (7.5 μL) of 480 μM ZnSO₄, containing 18.4 U nuclease S1 (*Aspergillus oryzae*, Sigma-Aldrich), 5 U Antarctic phosphatase (NEB) and specific amounts of labelled internal standards ([D₃]-m⁶₂A, [D₃]-m⁶A) were added, and the mixture was incubated at 37 °C for 2 h. After addition of 7.5 μL of a 520 μM [Na]₂-EDTA solution, containing 0.15 U snake venom phosphodiesterase I (*Crotalus adamanteus*, USB corporation), the sample was incubated for 2 h at 37 °C and then stored at -20 °C. Prior to UHPLC-MS/MS analysis, samples were filtered by using an AcroPrep Advance 96 filter plate 0.2 μm Supor (Pall Life Sciences).

UHPLC/MS-MS analysis of RNA samples.

Quantitative UHPLC-MS/MS analysis of digested RNA samples was performed using an Agilent 1290 UHPLC system equipped with a UV detector and an Agilent 6490 triple quadrupole mass spectrometer. For exact quantification of m⁶A and m⁶₂A internal quantification with the stable isotope dilution technique was developed (Supplementary Fig. 2 and Supplementary Tab. 2 and 3). The canonical nucleosides (C, U, G, A) were measured and quantified with the UV-detector (Supplementary Tab. 3). The source-dependent parameters were as follows: gas temperature 80 °C, gas flow 15 L/min (N₂), nebulizer 30 psi, sheath gas heater 275 °C, sheath gas flow 11 L/min (N₂), capillary voltage 2,500 V in the positive ion mode and nozzle voltage 500 V. The fragmentor voltage was 380 V. Delta EMV was set to 500 (positive mode). Compound-dependent parameters are summarized in Supplementary Table 1. Chromatography was performed by a Poroshell 120 SB-C8 column (Agilent, 2.7 μm, 2.1 mm × 150 mm).

at 35 °C using a gradient of water and MeCN, each containing 0.0075% (v/v) formic acid, at a flow rate of 0.35 mL/min: 0 →6 min; 0 →6.0% (v/v) MeCN; 6 →7 min; 6.0 →40% MeCN; 7 →10.8 min; 40 →80% MeCN; 10.8 →12.50 min; 80% MeCN; 12.50 →13.50 min; 80 →0% MeCN; 13.5 →15 min; 0% MeCN. The effluent up to 1.1 min and after 14 min was diverted to waste by a Valco valve. The autosampler was cooled to 4 °C. The injection volume was amounted to 39 µL.

Determination of extinction coefficients

For the following absolute internal quantification of nucleosides the exact concentration of the isotopically labeled internal standards needs to be known. In order to measure the concentrations one needs to determine the extinction coefficients first. For this determination, m^6A and m^6_2A were diluted in D_2O as well as the reference nucleoside U. The absorbance of all three solutions were measured as technical quintuplets on the *NanoDrop* 1000 Spectrophotometer for the λ_{max} of each nucleoside. Subsequently, a 1:1 mixture of m^6A and U as well as m^6_2A and U was prepared, each mixture as a technical triplicate. Following 1H -NMR measurements of these mixtures resulted in specific ratios of the amount n (m^6A) to the amount n (U), respectively n (m^6_2A) to n (U). 1H -NMR spectra were recorded on *Bruker* Avance III HD 800 MHz. By applying the Beer-Lambert law one can calculate the actual concentration c (U) solution with the literature known extinction coefficient (ϵ_λ). With the ratio of n (U) to n (m^6A), respectively n (U) to n (m^6_2A), one can calculate the concentrations c (m^6A) and c (m^6_2A) in each solution. Applying again the Beer-Lambert law results in the desired extinction coefficients of the two nucleosides ϵ_λ (m^6A) and ϵ_λ (m^6_2A) at a specific wavelength λ (see Supplementary Table 1).

Quantification of nucleosides

In order to obtain the internal calibration curves for exact quantification, each standard, namely $[D_3]$ - m^6A and $[D_3]$ - m^6_2A was analyzed in comparison to the corresponding non-labelled nucleoside with constant concentration. Technical triplicates were measured and the linear regression was applied using Origin® 6.0 (Microcal™). Therefore, the ratio of the area under the curve of unlabeled nucleoside (A) to the labelled standard (A*) was plotted against the ratio of the amount of unlabeled nucleoside (n) to the labelled one (n*) (see Supplementary Fig. 2). In order to obtain the UV calibration curves for quantification of the canonical nucleosides, namely C, U, G and A were measured in a concentration series. For every nucleoside two different solutions were diluted with a dilution factor of two in order to get five different concentrations for every nucleoside, which were measured as technical triplicates. The UV area under the curve (A) was plotted against the amount of the nucleoside (n) (see Supplementary Tab. 3). Acceptable precision (< 20% relative s.d. within each triplicate) and accuracy (80-120%) was achieved for all calibration curves. The accuracy is calculated as the ratio of the measured to the calculated ratios of the areas (A/A^*) under the curves in percent. The ratios of the areas (A/A^*) can be calculated by using the linear equations for the corresponding ratio of amount

LC-MS analysis of ORN-4

relative abundance

100
80
60
40
20
0

0 2 4 6 8 10 12 14 16 18 20

[min]

ORN 4 [855,8195]
-1Me [851,4818]
-2Me [846,4753]
-2Me [846,4745]
+ ORN 2 spike in
-3Me [841,8038]
-4Me [837,4655]

ORN 4
-2Me -1Me
 m^6A m^6A A

Supplementary Figure 3 HPL chromatogram of the incubation of ORN-4 with 16 μM ALKBH5₆₆₋₂₉₂ (black line) and with a spike in of ORN-2 (red line). The measured mass of the different species is shown in square brackets ($[\text{M}+3\text{H}]^{+3}$)

5. Supplementary Tables

Supplementary Table 1 Determined extinction coefficients for m^6A and m^6_2A .

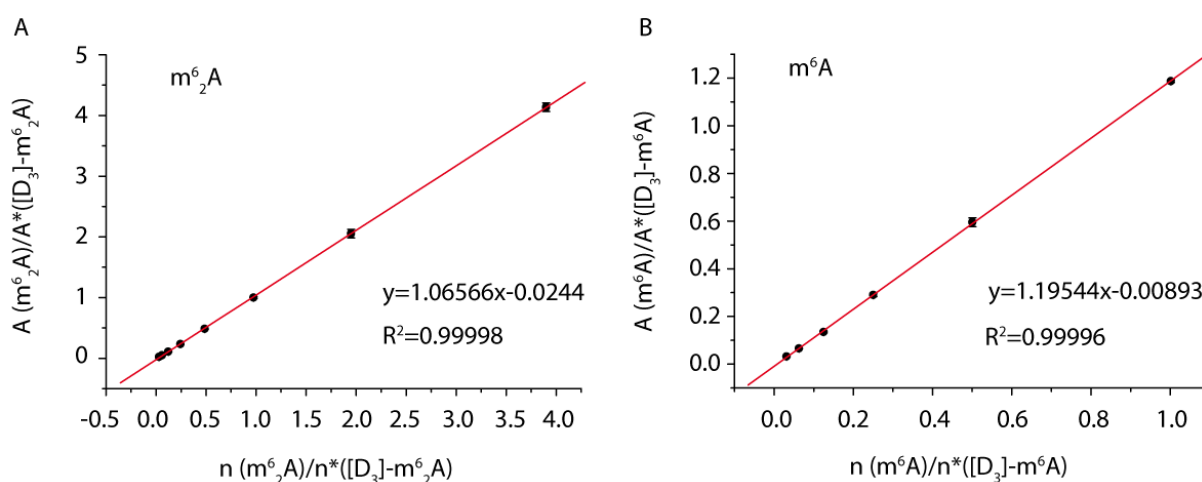
compound	$\epsilon_{\lambda_{\max}} (\lambda_{\max}) [L \cdot mol^{-1} \cdot cm^{-1}]$	$\epsilon_{\lambda=260nm} [L \cdot mol^{-1} \cdot cm^{-1}]$
m^6A	15474 (267 nm)	-
m^6_2A	19047 (274 nm)	12623

Supplementary Table 2 Compound-dependent LC-MS/MS-parameters used for the analysis of RNA. CE: collision energy, CAV: collision cell accelerator voltage, EMV: electron multiplier voltage. The nucleosides were analyzed in the positive ($[M+H]^+$ species) ion selected reaction monitoring mode (SRM).

compound	Precursor ion (m/z)	MS1 Resolution	Product Ion (m/z)	MS2 Resolution	Dwell time [ms]	CE (V)	CAV (V)	Polarity
Time segment 2.4-14 min								
$[D_3]-m^6_2A$	299.15	Wide	167.11	Wide	60	2	5	Positive
m^6_2A	296.14	Wide	164.11	Wide	60	2	5	Positive
$[D_3]-m^6A$	285.14	Wide	153.10	Wide	60	2	5	Positive
m^6A	282.12	Wide	150.08	Wide	60	2	5	Positive

Supplementary Table 3 Compound-dependent LC-MS/MS ranges of the corresponding linear equations.

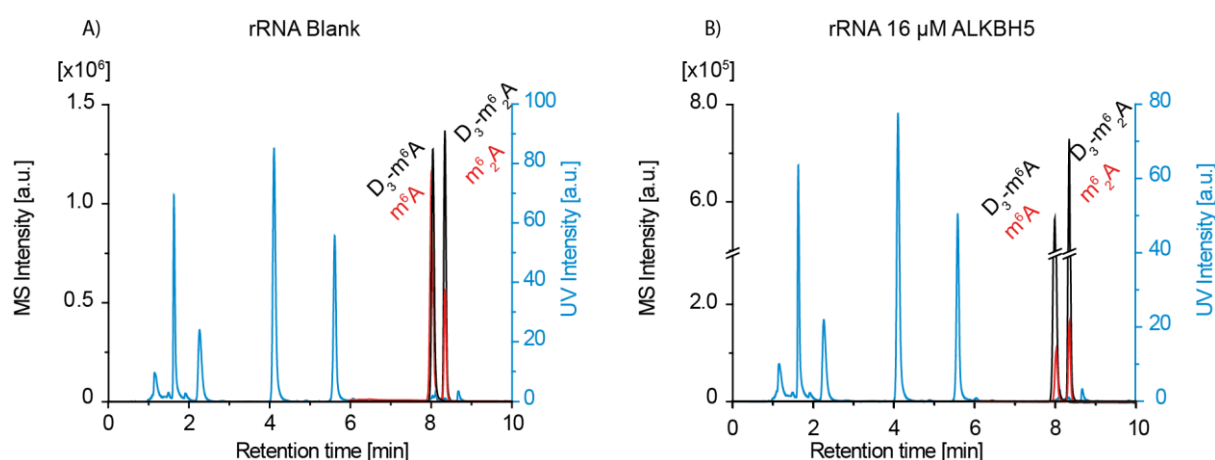
compound	n (ULOQ) [pmol]	n (LLOQ) [pmol]	A/A* (ULOQ)	A/A* (LLOQ)
m^6_2A	3.104	0.049	4.13	0.040
m^6A	1.709	0.053	1.18	0.029



Supplementary Figure 4 Internal calibration curves for the exact quantification of m^6_2A (A) and m^6A (B) with the corresponding linear equation and coefficient of determination.

Supplementary Table 4 Compound-dependent UV ranges and the corresponding linear equations of the canonical nucleosides.

compound	linear regression	n (ULOQ) [pmol]	n (LLOQ) [pmol]
C	$y = 0.28955x + 0.17515$	791	24.7
U	$y = 0.45953x - 0.34086$	792	24.7
G	$y = 0.56985x - 0.91256$	792	24.7
A	$y = 0.62408x - 0.10334$	659	20.6



Supplementary Figure 5 UV Chromatogram (blue) and MS trace (red: natural modification; black: isotopically labelled standard) of rRNA are shown, upon incubation with 16 μM ALKBH5₆₆₋₂₉₂ (B) and without ALKBH5₆₆₋₂₉₂ (A).

6. Literature

1. R. R. Edupuganti, S. Geiger, R. G. H. Lindeboom, H. Shi, P. J. Hsu, Z. Lu, S.-Y. Wang, M. P. A. Baltissen, P. W. T. C. Jansen, M. Rossa, M. Müller, H. G. Stunnenberg, C. He, T. Carell and M. Vermeulen, *Nat. Struct. Mol. Biol.*, 2017, **24**, 870-878.
2. C. Höbartner, C. Kreutz, E. Flecker, E. Ottenschläger, W. Pils, K. Grubmayr and R. Micura, *Monatsh. Chem.*, 2003, **134**, 851-873.

8. Experimenteller Teil

8.1. Chemische Synthesen

8.1.1. Allgemeine Arbeitsmethoden und Materialien

Alle Reaktionen wurden, wenn nicht anders angegeben, unter inerter Atmosphäre (N_2 oder Ar) und unter Verwendung der Standard-Schlencktechnik durchgeführt. Die verwendeten Glasgeräte wurden mehrfach bei 120 °C getrocknet und unter Vakuum abgekühlt. Chemikalien und Lösungsmittel wurden von den Firmen *ABCR*, *Acros*, *Fluka*, *Iris Biothech*, *Sigma-Aldrich*, *Carl Roth* und *TCI* in den gängigen Qualitätsstufen *puriss*, *p.a.* und *purum* bezogen und ohne weitere Aufreinigung verwendet. Für die Synthesen wurde, wenn nicht anders angegeben, über Molekularsieb getrocknetes Lösungsmittel verwendet und mit Einmalspritzen (Plastik, *Norm-Ject*) überführt. Für Extraktionen und säulenchromatographische Aufreinigungen wurde Lösungsmittel technischer Qualität genutzt und vor Gebrauch destilliert. Reaktionsmischungen wurden mit Magnetrührern *MR 3001 K* und *MR Hei-Tec* von der Firma *Heidolph* gerührt. Die Benennung der Moleküle in Abschnitt 9.1.2 erfolgte nicht in jedem Fall nach IUPAC-Regeln. Für eine bessere Übersichtlichkeit wurde in manchen Fällen eine kürzere Schreibweise gewählt, die nicht der IUPAC-Nomenklatur entspricht.

Dünnschichtchromatographie

Um den Fortschritt von Reaktionen und das Trenn- und Elutions-Verhalten bei der *flash*-Säulenchromatographie zu überprüfen wurden *Kieselgel 60 F254* Aluminiumplatten (*Merck*) verwendet. Die Visualisierung wurde durch UV Licht (254 nm oder 366 nm) und/oder einer Cerium Ammonium Molybdate Lösung (CAM, 10 g $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$, 2.0 g $Ce(SO_4)_2 \cdot 4H_2O$, 20 mL konz. H_2SO_4 in 180 mL ddH₂O), einer Vanilin Lösung (15 g Vanillin in 250 mL EtOH und 2.5 mL konz. H_2SO_4) oder einer Kaliumpermanganat Lösung (1.5 g $KMnO_4$, 10 g K_2CO_3 , 125 mg NaOH in 200 mL ddH₂O) erreicht.

Flash Säulenchromatographie

Für die säulenchromatografische Auftrennung wurde *Geduran® Si 60* Kieselgel (Normalphase, Korngröße: 0.040 – 0.063 mm) der Firma *Merck* genutzt. Die Auftrennung wurde unter leichtem Stickstoff-Überdruck durchgeführt. Bei Phosphoramiditen wurde Argon an Stelle von Stickstoff verwendet.

Rotationsverdampfer und Hochvakuum

Zur *in vacuo* Evaporieren der Lösungsmittel wurde ein *LABOROTA 4001-efficient* Rotationsverdampfer (*Heidolph*) in Verbindung mit einer *PC 3001 VARIO^{Pro}* Vakuumpumpe (*Vacuubrand*) verwendet. Das Hochvakuum wurde mit einer *RE 2.5* Drehschieberpumpe von der Firma *Vacuubrand* erreicht.

Kernspinresonanzspektroskopie (NMR-Spektroskopie) und IR

^1H , ^{13}C und ^{31}P NMR Spektren wurden in deuterierten Lösungsmitteln auf folgenden Spektrometern gemessen: *Avance III HD* (Bruker), *Ascend 400* (Bruker), *AMX 600* (Bruker), *XL 400* (Varian), *Inova 400* (Varian). Messungen auf 600 und 800 MHz und 2D Spektren (COSY, NOESY, HSQC, HMBC) wurden von der Analytik-Abteilung des Departments Chemie (Haus F) der Ludwig-Maximilians-Universität München durchgeführt. Tetramethylilan (Bei ^1H und ^{13}C Messungen) und H_3PO_4 (85 %, ^{31}P Messungen) dienten als externe Standards. Als interne Referenz dienten die deuterierten Lösungsmittel (nicht deuterierte Verunreinigung dieser) und deren Literaturbekannte Verschiebung.^[240] Die chemischen Verschiebungen (δ) werden in *parts per million* (ppm) angegeben und die Kopplungskonstanten (J) werden in Hertz (Hz) angegeben. Die Multiplizitäten der Signale werden als s (Singulet), d (Duplet), t (Triplet), q (Quartet), m (Multiplet) bzw. deren Kombinationen angegeben. IR Spektren wurden auf einem *Spectrum BX FT-IR* Instrument (*Perkin Elmer*) mit einem *DuraSamplIR II* Messkopfheit (*Smiths Detection*) aufgenommen und die Wellenzahl ($\tilde{\nu}$) werden in cm^{-1} angegeben. Die Intensitäten der Signale werden als schwach (w), medium (m) und stark (s) bezeichnet.

Hochauflösende Massenspektroskopie (ESI und EI) und LC-MS

Massenspektren zur Reaktionskontrolle wurden an einer *Dionex micro UHPLC* (LC, *Thermo Scientific*) gekoppelt mit einem *MSQ Single-Quadrupol* (MS, *Thermo Scientific*) aufgenommen. Hochauflösende Massenspektrometrie (HRMS) wurde von der Analytikabteilung des *Departments Chemie* der *Ludwig-Maximilians-Universität München* durchgeführt. Messungen mit der *electrospray ionization* (HRMS-ESI) Methode wurden an einer *Thermo Finnigan LTQ-FT* (*Thermo Scientific*) durchgeführt und Messungen mit *electron impact ionization* (HRMS-EI) Methode wurden an einer *Thermo Finnigan MAT 95* (*Thermo Scientific*) durchgeführt.

Reversed-Phase HPLC

Analytische *Reversed Phase high performance liquid chromatography* (RP-HPLC) Aufreinigungen wurde an einem *Waters Alliance*[®] 2695 bzw. e2695 *Separation Module* mit einem 2996 *Photodiode Array Detector* bzw. 2489 *UV/Visible Detector* und mit einem Fluss von 0.5 mL/min durchgeführt. Für die analytische Auftrennung wurden die Säulen EC 250/4 *Nucleodur* 100-3 C_{18} ec (Oligonukleotide) und EC 250/4 *Nucleosil* 120-3 C_{18} (*small molecules*), von *Machery & Nagel* benutzt. Semipräparative RP-HPLC Aufreinigungen wurden an einem *Waters 1525 Binary HPLC Pump* mit einem 2489 *UV/Visible Detector* und mit einem Fluss von 5 mL/min durchgeführt. Für die semipräparative Auftrennung wurden die Säulen VP 250/4 *Nucleodur* 100-5 C_{18} ec (Oligonukleotide) und VP 250/10 *Nucleosil* 100-5 C_{18} ec (*small molecules*), von *Machery & Nagel* benutzt. Zur Herstellung der Laufmittel und davon gepufferte Lösungen wurden ausschließlich Acetonitril in HPLC Qualität (HiPerSolv CHROMANORM, VWR) und frisch hergestelltes Reinstwasser (Leitfähigkeit $\cong 18.2 \text{ M}\Omega \times \text{cm}$) benutzt. Die benutzten Puffersysteme können aus der Tabelle (11) entnommen werden. Proben

für die RP-HPLC wurden vorher mit einem 0.2 µm GHP Spritzenfilter (13 mm) von *Acrodisc* gefiltert, um unlösliche Partikel zu entfernen.

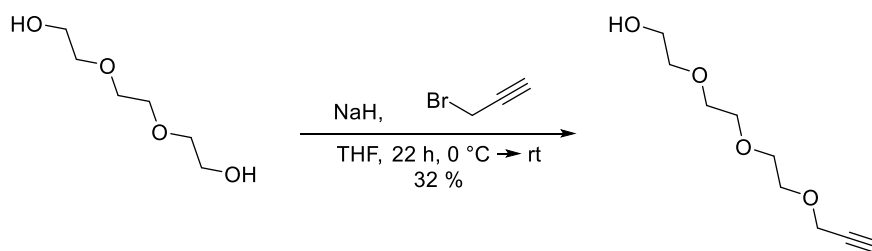
Tabelle 11: HPLC Puffersysteme, die in dieser Dissertation benutzt wurden.

Puffersystem	Zusammensetzung
HNEt ₃ OAc bei pH = 7.0	Puffer A = 0.1 M Triethylammoniumacetat in H ₂ O Puffer B = 0.1 M Triethylammoniumacetat in 80 % (v/v) Acetonitril
NH ₄ COOH bei pH = 4.3	Puffer A = 25 mM Ammoniumformiat in H ₂ O Puffer B = 25 mM Ammoniumformiat in 80 % (v/v) Acetonitril

8.1.2. Durchführung der Synthesen

8.1.2.1. Synthese der Haptene

O-Propargyl-triethylenglycol (2)



Triethylenglycol (**1**) (7.50 g, 50.0 mmol, 1.0 Äq) wurde in THF (37.5 mL) gelöst und unter Kühlung im Eisbad portionsweise Natriumhydrid (60 % in Mineralöl, 1.30 g, 32.5 mmol, 0.65 Äq) hinzugegeben. Die Reaktion wurde für 30 min bei 0 °C gerührt und anschließend innerhalb von 10 min 3-Brompropin (2.69 mL, 25.0 mmol, 0.5 Äq) hinzutropft. Die Lösung wurde für 2 h bei 0 °C weitergerührt und weitere 20 h bei Raumtemperatur, um die Reaktion zu vervollständigen. Der Mischung wurden 40 mL Wasser hinzugefügt und fünf Mal mit DCM (100 mL) extrahiert. Die vereinigte organische Phase wurde über Na₂SO₄ getrocknet und evaporiert. Das erhaltene Öl wurde einer säulenchromatographischen Aufreinigung unterzogen (isokratisch; *iso*Hexan/EtOAc 2:3) und das Produkt als gelblich-farbloses Öl zu erhalten (2.97 g, 15.8 mmol, 32 %).

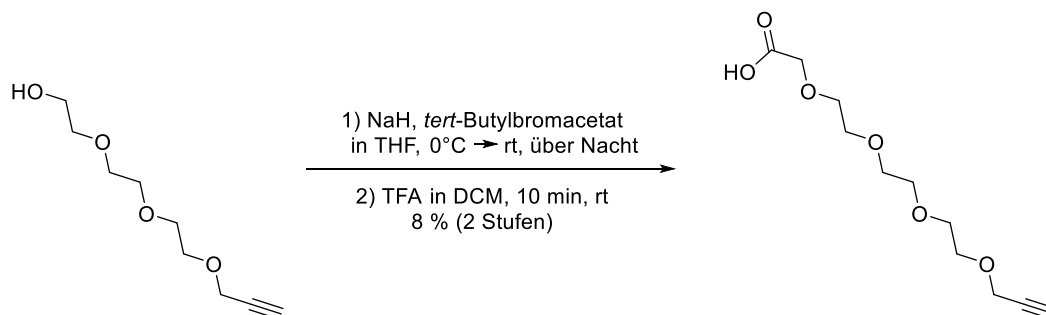
¹H NMR (599 MHz, CDCl₃) δ = 4.20 (d, *J* = 2.4 Hz, 2H, CH₂-C-CH_{Alkin}), 3.73 – 3.64 (m, 10H, CH₂), 3.60 (t, *J* = 4.3 Hz, 2H, CH₂), 2.43 (t, *J* = 2.4 Hz, 1H, CH_{Alkin}).

¹³C NMR (151 MHz, CDCl₃) δ = 79.7 (C-CH_{Alkin}), 74.7 (C-CH_{Alkin}), 72.6 (CH₂), 70.8 (CH₂), 70.5 (CH₂), 70.5 (CH₂), 69.2 (CH₂-O-CH₂-C-CH_{Alkin}), 61.9 (CH₂), 58.5 (CH₂-C-CH_{Alkin}).

HRMS-ESI (*m/z*): berechnet für C₉H₁₇O₄⁺ [M+H]⁺: 189.1121; gefunden: 189.1123

IR (ATR, neat): ν = 3436 (w), 3253 (w), 2870 (m), 2114 (w), 1734 (w), 1636 (w), 1456 (w), 1350 (m), 1288 (w), 1247 (w), 1091 (s), 1032 (s), 921 (m), 886 (m), 840 (m)

3,6,9,12-Tetraoxapentadec-14-insäure (3)



2 (2.90 g, 15.4 mmol, 1.0 Äq) wurde in THF (30 mL) gelöst und unter Kühlung im Eisbad portionsweise Natriumhydrid (60 % in Mineralöl, 0.86 g, 21.6 mmol, 1.4 Äq) hinzugegeben. Die Reaktion wurde für 30 min bei 0 °C gerührt und danach *tert*-Butyl Bromacetat (4.54 mL, 30.8 mmol, 2.0 Äq) langsam hinzugegeben. Die Lösung wurde für 1 h bei 0 °C gerührt und danach über Nacht bei Raumtemperatur weiter gerührt. Die Mischung wurde vorsichtig mit Wasser (40 mL) versetzt und fünf Mal mit DCM (100 mL) extrahiert. Die vereinigte organische Phase wurde über Na₂SO₄ getrocknet und das Lösungsmittel *in vacuo* entfernt. Das erhaltene Öl wurde einer säulenchromatographischen Aufreinigung unterzogen (gradient, DCM/MeOH, 19:1 → 9:1). Das erhaltene gelblich farblose Öl wurde in DCM (4 mL) gelöst und mit TFA (20 % in DCM, 1 mL) versetzt. Die Reaktion wurde für 10 min bei Raumtemperatur gerührt und das Lösungsmittel *in vacuo* entfernt. Das Produkt wurde fünf Mal mit DCM Coevaporiert, einer säulenchromatographischen Aufreinigung unterzogen (isokratisch; DCM/MeOH 9:1) und als farbloses Öl erhalten (0.30 g, 1.22 mmol, 8 % über 2 Stufen)

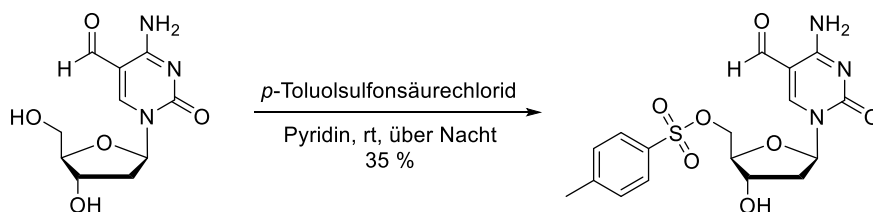
¹H NMR (400 MHz, CDCl₃) δ = 8.26 (s, b, 1H, COOH), 4.14 (d, J = 2.4 Hz, 2H, CH₂-C-CH_{Alkin}), 4.07 (s, 2H; COOH-CH₂), 3.70 – 3.59 (m, 12H, CH₂), 2.42 (t, J = 2.4 Hz, 1H, CH_{Alkin}).

¹³C NMR (101 MHz, CDCl₃) δ = 173.6 (COOH), 79.5 (CH₂-C-CH_{Alkin}), 74.8 (CH₂-C-CH_{Alkin}), 70.6 (CH₂), 70.3 (CH₂), 70.2 (CH₂), 70.2 (CH₂), 70.1 (CH₂), 68.9 (COOH-CH₂), 68.9 (CH₂), 58.3 (CH₂-C-CH_{Alkin}).

HRMS-ESI (m/z): berechnet für C₁₁H₁₇O₆[−] [M-H][−] : 245.1031; gefunden: 245.1025

IR (ATR, neat): ν = 3252 (w), 2872 (w), 1736 (m), 1456 (w), 1429 (w), 1350 (w), 1287 (w), 1245 (m), 1200 (m), 1092 (s), 1029 (s), 944 (m), 921 (m), 842 (m), 798 (w), 718 (m), 668 (s)

5'-O-(*p*-Toluolsulfonsäure)-5-formyl-2'-desoxycytidin (5)



5-Formylcytidin (150 mg, 587 μmol , 1.0 Äq) wurde in Pyridin (3 mL) gelöst und unter Kühlung im Eisbad *p*-Toluolsulfonylchlorid (134 mg, 705 μmol , 1.2 Äq), gelöst in Pyridin (1 mL), über 5 min hinzugeotropft. Die Reaktion wurde über Nacht bei Raumtemperatur gerührt, das Lösungsmittel *in vacuo* entfernt und einer säulenchromatographischen Aufreinigung unterzogen (Gradient; DCM/MeOH 49:1 \rightarrow 19:1). Das Produkt wurde als weißes Pulver erhalten (84.0 mg, 205 μmol , 35 %).

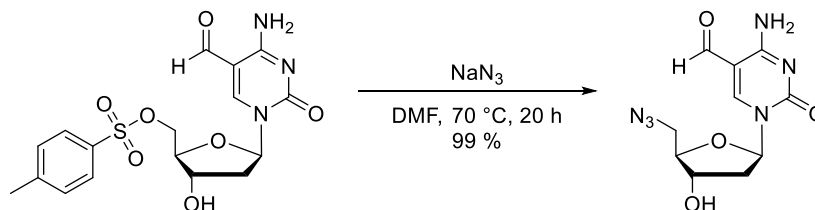
^1H NMR (599 MHz, MeOD) δ = 9.51 (s, 1H, $\text{H}_{\text{Aldehyd}}$), 8.60 (s, 1H, $\text{H}_{5, \text{Base}}$), 7.81 (d, J = 8.3 Hz, 2H, $\text{SO}_3\text{-C-CH}_{\text{ar}}$), 7.45 (d, J = 8.2 Hz, 2H, $\text{CH}_3\text{-C-CH}_{\text{ar}}$), 6.16 (t, J = 6.1 Hz, 1H, $\text{H}_{1'}$), 4.43 (dd, J = 11.2, 2.8 Hz, 1H, $\text{H}_{4'}$), 4.32 – 4.27 (m, 2H, $\text{H}_{5'}$), 4.15 (q, J = 3.9 Hz, 1H, $\text{H}_{3'}$), 2.52 (ddd, J = 13.9, 6.4, 4.7 Hz, 1H, $\text{H}_{2'}$), 2.46 (s, 3H, CH_3) 2.22 (dt, J = 13.9, 6.2 Hz, 1H, $\text{H}_{2'}$).

^{13}C NMR (151 MHz, MeOD) δ = 190.0 ($\text{C}_{\text{Aldehyd}}$), 164.3 ($\text{NH}_2\text{-C}$), 155.7 (O=C), 154.8 ($\text{C}_{5, \text{Base}}$), 147.1 ($\text{CH}_3\text{-C-CH}_{\text{ar}}$), 133.9 ($\text{SO}_3\text{-C}$), 131.3 ($\text{CH}_3\text{-C-CH}_{\text{ar}}$), 129.0 ($\text{SO}_3\text{-C-CH}_{\text{ar}}$), 106.8 ($\text{C}_{\text{Aldehyd-C}}$), 89.2 ($\text{C}_{1'}$), 86.2 ($\text{C}_{3'}$), 71.3 ($\text{C}_{5'}$), 70.2 ($\text{C}_{4'}$), 42.3 ($\text{C}_{2'}$), 21.6 (CH_3).

HRMS-ESI (m/z): berechnet für $\text{C}_{17}\text{H}_{20}\text{N}_3\text{O}_7\text{S}^+$ [$\text{M}+\text{H}$] $^+$: 410.1016; gefunden: 410.1015

IR (ATR, neat): ν = 3346 (w), 3257 (w), 1682 (w), 1633 (s), 1503 (m), 1420 (w), 1335 (m), 1303 (w), 1232 (m), 1188 (m), 1170 (s), 1107 (m), 1090 (m), 1063 (m), 960 (m), 923 (w), 905 (m), 876 (w), 810 (m), 784 (s), 693 (s)

5'-Azid-5-formyl-2'-desoxycytidin (6)



5 (57 mg, 139 μ mol, 1.0 Äq) wurden in DMF (4 mL) gelöst und Natriumazid (11 mg, 169 μ mol, 1.2 Äq) hinzugegeben. Die Reaktion wurde für 20 h bei 70 °C gerührt, das Lösungsmittel *in vacuo* entfernt und einer säulenchromatographischen Aufreinigung unterzogen (isokratisch; DCM/MeOH 9:1). Das Produkt wurde als weißes Pulver erhalten (39 mg, 139 μ mol, 99 %)

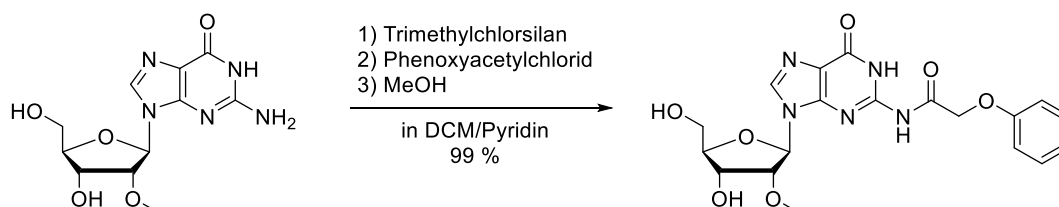
^1H NMR (400 MHz, MeOD) δ = 9.57 (s, 1H, $\text{H}_{\text{Aldehyd}}$), 8.66 (s, 1H, H_6), 6.19 (t, J = 6.3 Hz, 1H, $\text{H}_{1'}$), 4.34 (dt, J = 6.6, 4.4 Hz, 1H, $\text{H}_{3'}$), 4.12 – 4.08 (m, 1H, $\text{H}_{4'}$), 3.79 – 3.65 (m, 2H, $\text{H}_{5'}$), 2.53 (ddd, J = 13.9, 6.4, 4.5 Hz, 1H, $\text{H}_{2'}$), 2.29 (dt, J = 13.9, 6.4 Hz, 1H, $\text{H}_{2'}$).

^{13}C NMR (101 MHz, MeOD) δ = 189.7 (C-Aldehyd), 164.2 (C_4), 155.7 (C_2), 155.1 (C_6), 106.9 (C_5), 88.9 ($\text{C}_{1'}$), 86.9 ($\text{C}_{4'}$), 72.1 ($\text{C}_{3'}$), 53.2 ($\text{C}_{5'}$), 41.9 ($\text{C}_{2'}$).

HRMS-ESI (m/z): berechnet für $\text{C}_{10}\text{H}_{13}\text{N}_6\text{O}_4^+$ [$\text{M}+\text{H}$] $^+$: 281.0993; gefunden: 281.0996

IR (ATR, neat): ν = 3326 (w), 2928 (w), 2097 (m), 1634 (s), 1504 (s), 1431 (m), 1419 (m), 1405 (w), 1302 (m), 1268 (m), 1232 (s), 1193 (m), 1174 (m), 1074 (s), 1005 (m), 973 (m), 960 (m), 930 (m), 865 (w), 833 (w), 804 (w), 781 (s), 764 (m), 690 (m), 667 (s)

2'-O-Methyl- N^2 -phenoxyacetylguanosin (8)



2'-O-Methylguanosin (**7**) (200 mg, 673 μ mol, 1.0 Äq) wurde in Pyridin (5 mL) und DCM (13 mL) gelöst und über 2 min unter Eiskühlung Trimethylchlorosilan (658 mg, 6.1 mmol, 9.0 Äq) hinzugegeben.

Die Mischung wurde 2h bei rt, zur vollständigen Lösung, gerührt und unter erneuter Eiskühlung Phenoxyacetylchlorid (102 μ L, 740 μ mol, 1.1 Äq) über 30 min hinzugegeben. Die Reaktionsmischung wurde für 3.5 h unter Eiskühlung gerührt und nach vollständiger Umsetzung mit MeOH (5 mL) versetzt. Die Lösung wurde über Nacht bei Raumtemperatur gerührt und das Lösungsmittel *in vacuo* entfernt. Der ölige Rückstand wurde in Toluol (10 mL) aufgenommen und wieder eingeeengt. Dies wurde noch zwei Mal mit ddH₂O (je 10 mL) wiederholt und der verbleibende Rückstand in ddH₂O (20 mL) aufgenommen und unter Eiskühlung vollständig ausgefällt. Das Präzipitat wurde gefiltert und getrocknet, um das Produkt (290 mg, 672 μ mol, 99 %) als weißes Pulver zu erhalten.

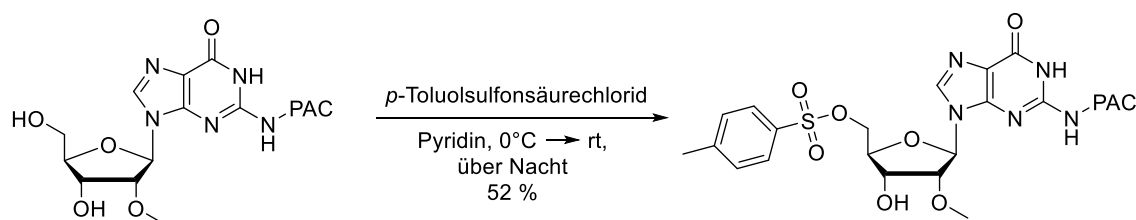
¹H NMR (400 MHz, DMSO-*d*₆) δ = 11.80 (s, 1H, NH), 11.79 (s, 1H, NH), 8.31 (s, 1H, CH_{Base}), 7.31 (ddd, *J* = 9.8, 6.8, 2.6 Hz, 2H, PAC_{meta}), 7.02 – 6.94 (m, 3H, PAC_{ortho/para}), 5.92 (d, *J* = 6.0 Hz, 1H, H_{1'}), 5.26 (d, *J* = 4.8 Hz, 1H, CH_{3'}-OH), 5.09 (s, 1H, CH_{5'}-OH), 4.87 (s, 2H, OCH₂), 4.32 (q, *J* = 4.3, 3.6 Hz, 1H, H_{3'}), 4.26 – 4.21 (m, 1H, H_{2'}), 3.94 (q, *J* = 3.9 Hz, 1H, H_{4'}), 3.68 – 3.53 (m, 2H, H_{5'}), 3.35 (s, 3H, CH₃).

¹³C NMR (101 MHz, DMSO-*d*₆) δ = 170.9 (O=C-NH), 157.6 (O-C-CH_{ar}), 154.9 (O=C_{6,Base}), 148.5 (C_{4,Base}), 147.4 (C_{2,Base}), 137.7 (C_{8,Base}), 129.5 (PAC_{meta}), 121.3 (PAC_{ortho/para}), 120.4 (C_{5,Base}), 114.6 (PAC_{ortho/para}), 86.1 (C_{4'}), 84.8 (C_{1'}), 82.9 (C_{2'}), 68.6 (C_{3'}), 66.2 (OCH₂), 61.1 (C_{5'}), 57.6 (CH₃).

HRMS-ESI (*m/z*): berechnet für C₁₉H₂₂N₅O₇⁺ [M+H]⁺: 432.1514; gefunden: 432.1514

IR (ATR, neat): ν = 3154 (w, sehr breit), 2360 (w), 1687 (s), 1611 (m), 1588 (w), 1557 (m), 1487 (m), 1432 (w), 1410 (w), 1361 (w), 1290 (w), 1262 (m), 1245 (m), 1209 (s), 1175 (w), 1128 (m), 1088 (m), 1051 (m), 990 (w), 975 (w), 947 (w), 845 (w), 780 (m), 749 (s), 712 (w), 688 (m), 673 (w)

5'-*O*-(*p*-Toluolsulfonsäure)-2'-*O*-methyl-*N*²-phenoxyacetylguanosin (**9**)



Das PAC-Gm **8** (300 mg, 0.7 mmol, 1.0 Äq) wurde in Pyridin (2 mL) gelöst und unter Eiskühlung in Pyridin (1 mL) gelöstes *p*-Toluolsulfonylchlorid (265 mg, 1.4 mmol, 2.0 Äq) über 10 min zur Lösung getropft. Die Lösung wurde für 12 h unter Eiskühlung gerührt und danach mit EtOAc (50 mL) verdünnt.

Die organische Phase wurde mit einer NaCl Lösung (50 % aq, 2 x 20 mL) gewaschen, über Na₂SO₄ getrocknet und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde einer säulenchromatographischen Aufreinigung unterzogen (isokratisch, DCM/MeOH, 97:3). Das Produkt (210 mg, 359 µmol, 52 %) wurde als leicht bräunlicher Feststoff erhalten.

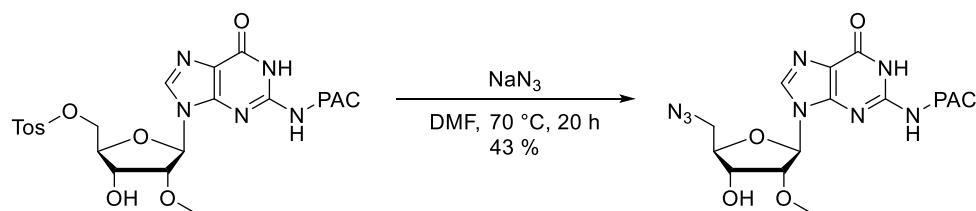
¹H NMR (599 MHz, MeOD) δ = 8.06 (s, 1H, CH_{8,Base}), 7.72 (d, *J* = 8.3 Hz, 2H, SO₃-C-CH_{ar}), 7.36 – 7.29 (m, 4H, PAC_{meta}/CH₃-C-CH_{ar}), 7.06 (d, *J* = 8.6 Hz, 2H, PAC_{ortho}), 7.03 (t, *J* = 7.4 Hz, 1H, PAC_{para}), 6.00 (d, *J* = 4.2 Hz, 1H, H_{1'}), 4.84 (s, 2H, OCH₂), 4.51 (t, *J* = 5.3 Hz, 1H, H_{3'}), 4.39 – 4.36 (m, 2H, H_{5'}), 4.33 (t, *J* = 4.7 Hz, 1H, H_{2'}), 4.20 (td, *J* = 4.9, 3.1 Hz, 1H, H_{4'}), 3.50 (s, 3H, OMe bei 2'), 2.41 (s, 3H, Tos-CH₃).

¹³C NMR (101 MHz, MeOD) δ = 172.7 (O=C-NH), 159.0 (O-C-CH_{ar}), 157.3 (O=C_{6,Base}), 148.6 (C), 146.7 (C), 140.0 (C_{8,Base}), 133.9 (C), 130.9 (PAC_{meta}/CH₃-C-CH_{ar}), 130.7 (PAC_{meta}/CH₃-C-CH_{ar}), 128.9 (SO₃-C-CH_{ar}), 123.0 (PAC_{para}), 115.9 (PAC_{ortho}), 88.5 (C_{1'}), 84.1 (C_{2'}), 83.5 (C_{4'}), 70.6 (C_{5'}), 70.4 (C_{3'}), 68.1 (OCH₂), 59.1 (OCH₃ bei 2'), 21.6 (Tos-CH₃).

HRMS-ESI (*m/z*): berechnet für C₂₆H₂₈N₅O₉S⁺ [M+H]⁺: 586.1602; gefunden: 586.1608

IR (ATR, neat): ν = 3192 (w), 2360 (w), 1684 (s), 1653 (w), 1597 (m), 1558 (s), 1540 (m), 1506 (w), 1489 (m), 1457 (m), 1404 (m), 1360 (m), 1221 (m), 1189 (m), 1173 (s), 1086 (m), 972 (m), 783 (s), 753 (s), 719 (m), 690 (s), 665 (s)

5'-Azid-2'-*O*-methyl-*N*²-phenoxyacetylguanosin (10)



5'-Tosyl-2PAC-2'Methylguanosin (60 mg, 102 µmol, 1.0 Äq) wurden in DMF (1.2 mL) gelöst und Natriumazid (37 mg, 570 µmol, 5.6 Äq) hinzugegeben. Die Reaktionsmischung wurde für 4.5 h bei 75 °C gerührt, anschließend mit Wasser (5 mL) verdünnt und mit EtOAc (5 mL) extrahiert. Die Organische Phase wurde 5 Mal mit einer NaCl Lösung (50 % aq., 15 mL) gewaschen, über Na₂SO₄ getrocknet und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde einer

säulenchromatographischen Aufreinigung unterzogen (Gradient, DCM/MeOH, 97:3 → 25:1). Es wurde das Produkt (20 mg, 43.2 μ mol, 43 %) als weißes Pulver erhalten.

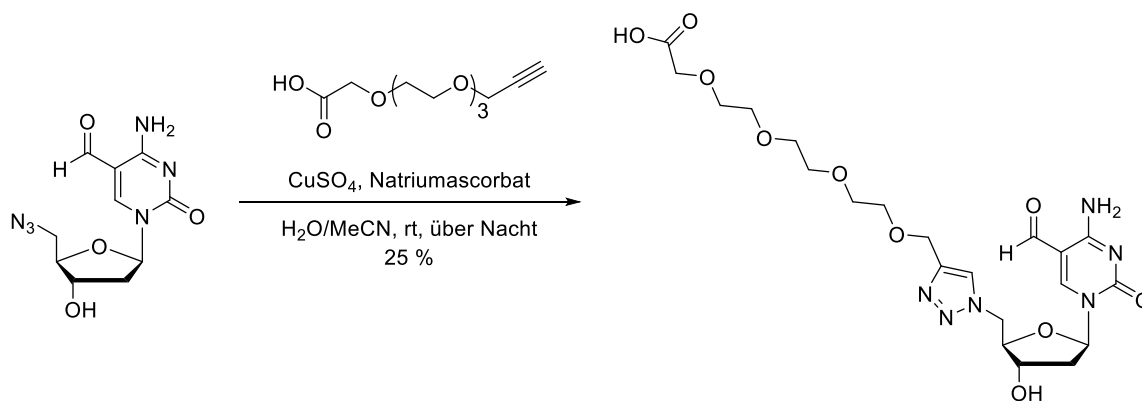
^1H NMR (599 MHz, MeCN- d_3) δ 7.95 (s, 1H, $\text{CH}_{8,\text{Base}}$), 7.36 – 7.32 (m, 2H, PAC_{meta}), 7.06 – 7.02 (m, 3H $\text{PAC}_{\text{ortho/para}}$), 5.90 (d, J = 4.7 Hz, 1H, $\text{H}_{1'}$), 4.79 (s, 2H, OCH_2), 4.43 (t, J = 5.0 Hz, 1H, $\text{H}_{3'}$), 4.36 (t, J = 5.0 Hz, 1H, $\text{H}_{2'}$), 4.09 (q, J = 5.2 Hz, 1H, $\text{H}_{4'}$), 3.61 (dd, J = 9.9, 4.7 Hz, 2H, $\text{H}_{5'}$), 3.43 (s, 3H, OMe bei 2').

^{13}C NMR (151 MHz, MeCN- d_3) δ = 172.0 ($\text{O}=\text{C}-\text{NH}$), 158.2 ($\text{O}-\text{C}-\text{CH}_{\text{ar}}$), 156.3 (C), 149.3 ($\text{C}_{4,\text{Base}}$), 148.2 (C), 139.0 ($\text{C}_{8,\text{Base}}$), 130.7 (PAC_{meta}), 123.0 (PAC_{para}), 122.6 ($\text{C}_{5,\text{Base}}$), 115.7 ($\text{PAC}_{\text{ortho}}$), 87.4 ($\text{C}_{1'}$), 84.2 ($\text{C}_{4'}$), 83.4 ($\text{C}_{2'}$), 71.0 ($\text{C}_{3'}$), 67.7 (OCH_2), 59.0 (OCH_3 bei 2'), 52.8 ($\text{C}_{5'}$)

HRMS-ESI (m/z): berechnet für $\text{C}_{19}\text{H}_{19}\text{N}_8\text{O}_6^-$ [$\text{M}-\text{H}$] $^-$: 455.1433; gefunden: 455.1432

IR (ATR, neat): ν = 3210 (w), 2933 (w), 2103 (m, Azid), 1689 (s), 1611 (s), 1598 (s), 1585 (s), 1562 (s), 1495 (m), 1406 (m), 1363 (w), 1291 (w), 1239 (m), 1214 (m), 1176 (m), 1130 (m), 1080 (m), 1042 (w), 995 (w), 832 (w), 785 (w), 755 (m), 692 (w)

5'-N-((1H-1,2,3-Triazol-4-yl)-3,6,9,12-tetraoxatridecansäure)-5-formyl-2'-desoxycytidin (11)



Das Azid-fdC **6** (13 mg, 46.4 μ mol, 1.0 Äq) wurde unter Argon in Wasser (1 mL) gelöst und der PEG3-Linker **3** (11 mg, 46.4 μ mol, 1.0 Äq, gelöst in 200 μ L MeCN) wurde hinzugegeben. Die Reaktionsmischung wurde 2 Mal mit Argon entgast („Freeze-Pump-Thaw“), CuSO_4 (7.4 mg, 46.4 μ mol, 1.0 Äq) dazugegeben und mit Argon ein weiteres Mal entgast. Die Cycloadditionsreaktion wurde durch Zugabe von Natriumascorbat (3.7 mg, 18.6 μ mol, 0.4 Äq) gestartet und erneut entgast. Die Reaktionsmischung wurde über Nacht bei Raumtemperatur gerührt und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde in ddH₂O (3 mL) aufgenommen und zweifach über semipräparative

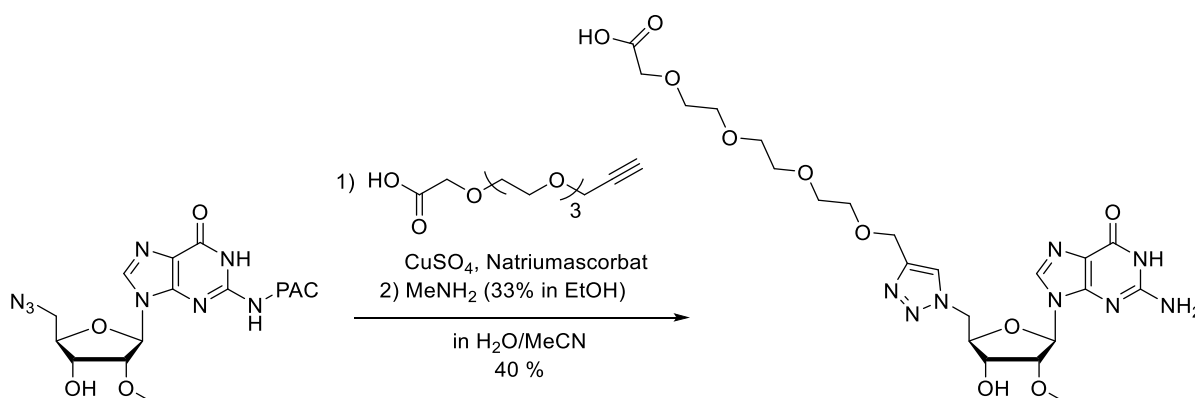
HPLC (jeweils Gradient über 45 min; Puffer A = 25 mM NH₄COOH in H₂O, pH = 4.3; Puffer B = 25 mM NH₄COOH in 80 % MeCN_{aq}, pH = 4.3; 0 → 35 % Puffer B) einer Aufreinigung unterzogen. Das Produkt wurde als weißer Schaum (6.0 mg, 11.4 μmol, 25 %) erhalten.

¹H NMR (599 MHz, D₂O) δ = 9.50 (s, 1H, **H**_{Aldehyd}), 8.50 (s, 1H, **NH**₂), 8.26 (s, 1H, **H**_{ar}), 8.11 (s, 1H, **H**_{Triazol}), 6.15 (t, *J* = 6.0 Hz, 1H, **H**_{1'}), 4.90 – 4.85 (m, 1H, **H**_{5'}), 4.83 – 4.76 (m, 1H, **H**_{5'}), 4.65 (s, 2H, **C**_{Triazol}-**CH**₂), 4.48 (q, *J* = 5.6 Hz, 1H, **H**_{4'}), 4.43 (q, *J* = 5.2 Hz, 1H, **H**_{3'}), 3.94 (s, 2H, **COOH-CH**₂), 3.73 – 3.63 (m, 12H, **CH**₂), 2.62 – 2.53 (m, 1H, **H**_{2'}), 2.36 – 2.27 (m, 1H, **H**_{2'}).

¹³C NMR (151 MHz, D₂O) δ = 190.1 (**C**_{Aldehyd}), 177.5 (**COOH**), 170.5 (**C=O**), 162.4 (**C**_{ar}), 154.9 (**C**_{ar}), 154.4 (**CH**_{ar}), 144.2 (**C**_{Triazol}), 126.1 (**CH**_{Triazol}), 87.5 (**C**_{1'}), 83.9 (**C**_{4'}), 70.2 (**C**_{3'}), 69.6 (**COOH-CH**₂), 69.5 (**CH**₂), 69.4 (**CH**₂), 69.4 (**CH**₂), 69.3 (**CH**₂), 68.9 (**CH**₂), 62.9 (**C**_{Triazol-CH}₂), 50.7 (**C**_{5'}), 38.4 (**C**_{2'}).

HRMS-ESI (*m/z*): berechnet für C₂₁H₂₉N₆O₁₀[−] [M-H][−]: 525.1951; gefunden: 525.1950

5'-N-((1H-1,2,3-Triazol-4-yl)-3,6,9,12-tetraoxatridecansäure)-2'-O-methyl-N²-phenoxyacetyl-guanosin (12)



10 (13.0 mg, 28.5 μmol, 1.0 Äq) wurde unter Argon in Wasser (1 mL) gelöst und der PEG3-Linker **3** (7.0 mg, 28.5 μmol, 1.0 Äq, gelöst in 200 μL MeCN) wurde hinzugegeben. Die Reaktionsmischung wurde 2 Mal mit Argon entgast („Freeze-Pump-Thaw“), CuSO₄ (4.6 mg, 28.5 μmol, 1.0 Äq) dazugegeben und mit Argon ein weiteres Mal entgast. Die Cycloadditionsreaktion wurde durch Zugabe von Natriumascorbat (2.3 mg, 11.4 μmol, 0.4 Äq) gestartet und die Lösung erneut entgast. Die Reaktionsmischung wurde über Nacht bei Raumtemperatur gerührt und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde über semipräparative HPLC (Gradient über 45 min; Puffer A = 25 mM NH₄COOH in H₂O, pH = 4.3; Puffer B = 25 mM NH₄COOH in 80 % MeCN_{aq}, pH = 4.3; 0 → 65 % Puffer B) einer Aufreinigung unterzogen. Die Elution wurde lyophilisiert und anschließend in MeNH₂ (33 % in EtOH, 1 mL, 8.5 mmol, 400 Äq) und ddH₂O (100 μL) aufgenommen und für 2 min

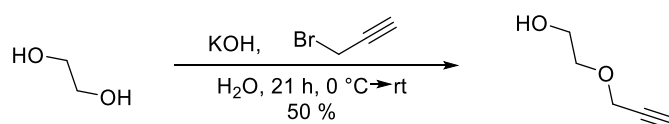
bei Raumtemperatur gerührt. Das Lösungsmittel wurde *in vacuo* entfernt. Der Rückstand wurde in ddH₂O (3 mL) aufgenommen und über semipräparative HPLC (Gradient über 45 min; Puffer A = 25 mM NH₄COOH in H₂O, pH = 4.3; Puffer B = 25 mM NH₄COOH in 80 % MeCN_{aq}, pH = 4.3; 5 → 15 % Puffer B) einer Aufreinigung unterzogen. Das Produkt wurde als weißer Feststoff (8.0 mg, 11.4 μmol, 40 %) erhalten.

¹H NMR (800 MHz, DMSO-*d*₆) δ = 8.36 (s, 1H, NH_{1,Base}, schwach), 7.88 (s, 1H, H_{Triazol}), 7.82 (s, 1H, CH_{8,Base}), 7.09 (s, 2H, NH₂), 5.79 (d, *J* = 6.4 Hz, 1H, H_{1'}), 4.84 (dd, *J* = 14.2, 8.9 Hz, 1H, H_{5'}), 4.72 (dd, *J* = 14.3, 3.8 Hz, 1H, H_{5'}), 4.51 – 4.42 (m, 3H, H_{2'} und C_{Triazol}-CH₂), 4.39 (dd, *J* = 4.6, 3.0 Hz, 1H, H_{3'}), 4.25 (dt, *J* = 8.8, 3.4 Hz, 1H, H_{4'}), 3.74 (s, 2H, COOH-CH₂), 3.55 – 3.45 (m, 12H, CH₂), 3.31 (s, 3H, O-CH₃).

¹³C NMR (201 MHz, DMSO-*d*₆) δ = 173.0 (COOH), 164.8 (C), 157.0 (C), 154.3 (C), 150.9 (C), 143.7 (C_{Triazol}), 136.2 (CH_{8,Base}), 124.7 (CH_{Triazol}), 117.2 (C), 86.4 (schwach, C_{1'}), 83.4 (schwach, C_{4'}), 80.8 (C_{2'}), 69.9 (schwach, COOH-CH₂), 69.7 (CH₂), 69.7 (CH₂), 69.6 (CH₂), 69.6 (CH₂), 69.3 (C_{3'}), 69.1 (CH₂), 69.0 (CH₂), 63.5 (C_{Triazol}-CH₂), 57.6 (O-CH₃), 51.4 (C_{5'}).

HRMS-ESI (*m/z*): berechnet für C₂₂H₃₁N₈O₁₀[−] [M-H][−]: 567.2169; gefunden: 567.2173

O-Propargyl-ethylenglycol (15) ^[195]



Ethylenglycol **14** (9.0 mL, 162 mmol, 4.3 Äq) wurde in Wasser (5.7 mL) gelöst und unter Eiskühlung KOH (3.6 g, 64 mmol, 1.7 Äq) hinzugegeben. Zur entstehenden Lösung wurde Propargylbromid (80 % in Toluol, 5.6 mL, 38 mmol, 1.0 Äq) über 30 min zu getropft. Die Lösung wurde für 21 h bei Raumtemperatur gerührt. Die Mischung wurde mit Wasser (60 mL) versetzt und mit DCM extrahiert (6 x 100 mL). Die vereinigten organischen Phasen wurden über Na₂SO₄ getrocknet und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde mittels Säulenchromatographie einer Aufreinigung unterzogen (isokratisch, iHex/EtOAc, 2:1), um das Produkt (1.9 g, 19.1 mmol, 50 %) als weißen Feststoff zu erhalten.

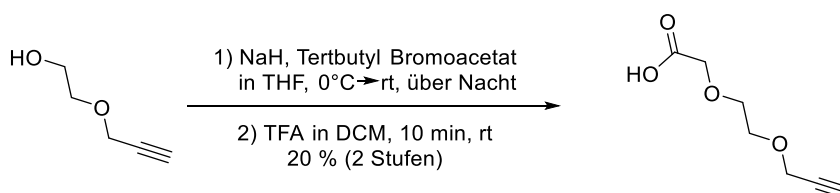
¹H NMR (400 MHz, CDCl₃) δ = 4.21 (d, *J* = 2.3 Hz, 2H, CH₂-C-CH_{alkin}), 3.80 – 3.74 (t, *J* = 4.2 Hz, 2H, OH-CH₂), 3.68 – 3.62 (t, *J* = 4.2 Hz, 2H, OH-CH₂-CH₂), 2.46 (t, *J* = 2.3 Hz, 1H, CH_{alkin}).

^{13}C NMR (101 MHz, CDCl_3) δ = 79.6 ($\text{C}-\text{CH}_{\text{Alkin}}$), 74.9 (CH_{Alkin}), 71.3 (CH_2), 61.9 (CH_2), 58.6 ($\text{CH}_2-\text{C}-\text{CH}_{\text{Alkin}}$).

HRMS-ESI (m/z): berechnet für $\text{C}_5\text{H}_7\text{O}_2^-$ $[\text{M}-\text{H}]^-$: 99.0452; gefunden: 99.0441

IR (ATR, neat): ν = 3393 (m), 3287 (m), 2932 (w), 2869 (w), 2118 (w), 1717 (w), 1641 (w), 1460 (w), 1444 (w), 1408 (w), 1354 (m), 1278 (w), 1219 (w), 1105 (s), 1066 (s), 1028 (s), 1006 (m), 953 (w), 919 (w), 889 (w), 826 (w)

2-(2-(Prop-2-in-1-yloxy)ethoxy)essigsäure (16) ^[195]



15 (1.53 g, 15.2 mmol, 1.0 Äq) wurde in THF (40 mL) gelöst und unter Eiskühlung Natriumhydrid (0.54 g, 22.5 mmol, 1.5 Äq) portionsweise hinzugegeben. Nachdem die Lösung Raumtemperatur erreicht hat, wurde die Lösung eine halbe Stunde bei Raumtemperatur gerührt. Unter erneuter Eiskühlung wurde Tert-butyl Bromacetat (3.2 mL, 22.5 mmol, 1.5 Äq) über 5 min zur Lösung zu getropft. Die Reaktionsmischung wurde für 20 h bei Raumtemperatur weiter gerührt. Das Lösungsmittel wurde *in vacuo* entfernt, der Überstand in Wasser (100 mL) aufgenommen und 3 Mal mit DCM (80 mL) extrahiert. Die vereinigte organische Phase wurde über Na_2SO_4 getrocknet und das Lösungsmittel *in vacuo* entfernt. Das erhaltene Öl wurde einer säulenchromatographischen Aufreinigung unterzogen (isokratisch, *i*Hex/EtOAc, 4:1). Das erhaltene gelbliche Öl wurde mit TFA (20% in DCM, 25 mL) versetzt und für 2 h bei Raumtemperatur gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand in DCM (20 mL) aufgenommen. Die organische Phase wurde 5 Mal mit Wasser (30 mL) gewaschen. Die organische Phase wurde über Na_2SO_4 getrocknet, das Lösungsmittel *in vacuo* entfernt und das Produkt als gelbliches Öl erhalten (0.48 g, 3.04 mmol, 20% über 2 Stufen).

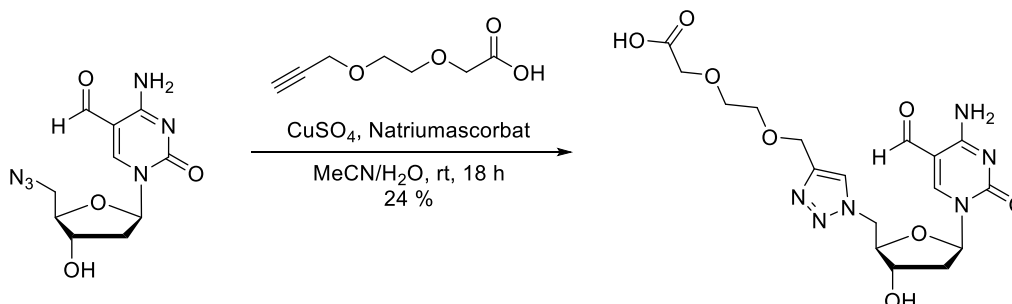
^1H NMR (599 MHz, CDCl_3) δ = 9.53 (s, 1H, COOH), 4.21 (d, J = 2.4 Hz, 2H, $\text{CH}_2-\text{C}-\text{CH}_{\text{alkin}}$), 4.19 (s, 2H, COOH- CH_2), 3.79 – 3.76 (m, 2H, CH_2), 3.76 – 3.72 (m, 2H, CH_2), 2.45 (t, J = 2.4 Hz, 1H).

^{13}C NMR (151 MHz, CDCl_3) δ = 174.7 (COOH), 79.2 ($\text{C}-\text{CH}_{\text{alkin}}$), 75.1 (CH_{alkin}), 71.0 (CH_2 bei 3.79 – 3.76), 69.0 (CH_2 bei 3.76 – 3.72), 68.4 (COOH- CH_2), 58.6 ($\text{CH}_2-\text{C}-\text{CH}_{\text{alkin}}$).

HRMS-ESI (m/z): berechnet für $C_7H_9O_4^-$ $[M-H]^-$: 157.0506; gefunden: 157.0507

IR (ATR, neat): ν = 3280 (w), 2920 (w), 2117 (w), 1727 (s), 1433 (w), 1354 (w), 1202 (m), 1141 (s), 1092 (s), 1026 (m), 920 (m), 880 (m), 840 (m)

5'-N-((1H-1,2,3-triazol-4-yl)-3,6,-dioxahheptansäure)-5-formyl-2'-desoxycytidin (17)



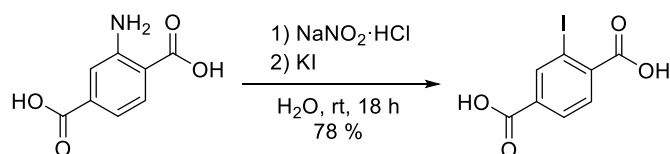
5'-Azid-5-formylcytidin (13 mg, 46.4 μ mol, 1.0 Äq) wurde unter Argon in Wasser (1 mL) gelöst und das *Linker*-Alkin **16** (7.3 mg, 46.4 μ mol, 1.0 Äq, gelöst in 200 μ L MeCN) wurde zur Mischung hinzugegeben. Die Mischung wurde 2 Mal mit Argon entgast („Freeze-Pump-Thaw“), $CuSO_4$ (7.4 mg, 46.4 μ mol, 1.0 Äq) dazugegeben und mit Argon ein weiteres Mal entgast. Die Cycloadditionsreaktion wurde durch Zugabe von Natriumascorbat (3.7 mg, 18.6 μ mol, 0.4 Äq) gestartet und erneut entgast. Die Reaktionsmischung wurde über Nacht bei Raumtemperatur gerührt und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde in ddH₂O (3 mL) aufgenommen und zweifach über semipräparative HPLC (jeweils Gradient über 45 min; Puffer A = 25 mM NH_4COOH in H₂O, pH = 4.3; Puffer B = 25 mM NH_4COOH in 80 % MeCN_{aq}, pH = 4.3; 0 \rightarrow 45min 20 % Puffer B) einer Aufreinigung unterzogen. Das Produkt wurde als weißer Schaum (4.8 mg, 11.0 μ mol, 24 %) erhalten.

¹H NMR (599 MHz, D₂O) δ = 9.50 (s, 1H, **H**_{Aldehyd}), 8.26 (s, 1H, **H**_{ar}), 8.17 (s, 1H, **H**_{Triazol}), 6.16 (t, J = 5.9 Hz, 1H, **H**_{1'}), 4.90 – 4.81 (m, 2H, **H**_{5'}), 4.65 (s, 2H, **CH**₂), 4.51 – 4.45 (m, 1H, **H**_{3'}), 4.47 – 4.41 (m, 1H, **H**_{4'}), 3.72 (m, 2H, **CH**₂), 3.56 – 3.50 (m, 2H, **CH**₂), 3.39 – 3.33 (m, 4H, **CH**₂), 2.61 – 2.55 (m, 1H, **H**_{2'}), 2.33 (m, 1H, **H**_{2'}).

¹³C NMR (151 MHz, D₂O) δ = 190.1 (**C**_{Aldehyd}), 162.5 (**COOH**), 155.0 (**C=O**), 154.4 (**CH**_{ar}), 105.7 (**CH**_{Triazol}), 87.5 (**C**_{1'}), 83.8 (**C**_{4'}), 70.2 (**C**_{3'}), 69.3 (**C**), 69.1 (**CH**₂), 63.9 (**C**), 63.0 (**CH**₂), 52.3 (**CH**₂), 52.0 (**C**), 50.7 (**C**_{5'}), 44.8 (**CH**₂), 38.3 (**C**_{2'}).

HRMS-ESI (m/z): berechnet für $C_{17}H_{21}N_6O_8^-$ $[M-H]^-$: 437.1426; gefunden: 437.1437

8.1.2.2. Synthese der wasserlöslichen IBX-Derivate

2-Iod-terephthalsäure (**21**) ^[229]

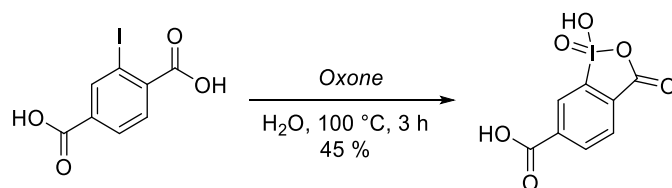
2-Amino-terephthalsäure (**20**) (1.0 g, 5.5 mmol, 1.0 Äq) wurde in HCl_{aq} (6 M, 30 mL) gelöst und unter Eiskühlung NaNO₂ (gelöst in 5 mL H₂O, 1.4 g, 20.3 mmol, 3.7 Äq) über 45 min zugegeben. Die Lösung wurde 30 min bei 0 °C weiter gerührt und vorsichtig in eine wässrige KI-Lösung (in 50 mL H₂O, 5.5 g, 33.0 mmol, 6.0 Äq) gegeben. Die Reaktionsmischung wurde für 18 h bei Raumtemperatur weitergerührt und mit NaHSO₃ versetzt, bis ein Farbumschwung zu erkennen war und ein gelbes Präzipitat ausfiel. Die Suspension wurde gefiltert und der Filterkuchen mit DCM/H₂O (1:1) gewaschen. Das Produkt (1.26 g, 16.6 mmol, 78 %) wurde als gelber Feststoff erhalten.

¹H NMR (400 MHz, Acetone-*d*₆) δ = 8.60 (d, *J* = 1.5 Hz, 1H, IC-CH_{ar}), 8.12 (dd, *J* = 8.0, 1.6 Hz, 1H, CH_{ar}-C-CH_{ar}), 7.94 (d, *J* = 8.1 Hz, 1H, IC-C-CH_{ar})

¹³C NMR (101 MHz, Acetone-*d*₆) δ = 167.6 (COOH), 165.5 (COOH), 142.6, 141.2, 134.7, 131.2, 130.0, 93.5 (IC)

HRMS-ESI (*m/z*): berechnet für C₈H₄IO₄⁻ [M-H]⁻ : 290.9160; gefunden: 290.9160

IR (ATR, neat): ν = 2814 (m), 2650 (m), 2536 (m), 1822 (w), 1684 (s), 1555 (w), 1479 (m), 1410 (m), 1374 (w), 1293 (s), 1256 (s), 1154 (w), 1136 (m), 1035 (m), 914 (m), 864 (w), 782 (m), 746 (s), 689 (w), 675 (w)

4-Carboxy-2-iod(V)oxybenzoesäure (**22**) ^[229]

2-Iod-Terephthalsäure (**21**) (150 mg, 0.5 mmol, 1.0 Äq) wurde in H₂O (4 mL) gelöst und OXONE[®] (753 mg, 2.4 mmol, 4.8 Äq) portionsweise zugegeben. Die Lösung wurde für 30 min bei 100 °C erhitzt

und ein weiteres Mal OXONE[®] hinzugegeben, bis sich das gebildete Präzipitat wieder gelöst hatte. Die Reaktionsmischung wurde heiß gefiltert und auf Raumtemperatur abgekühlt. Das Präzipitat wurde abgefiltert und mit H₂O gewaschen. Das Präzipitat wurde aus H₂O umkristallisiert, um das Produkt (64 mg, 0.22 mmol, 45 %) als farblosen Feststoff zu erhalten.

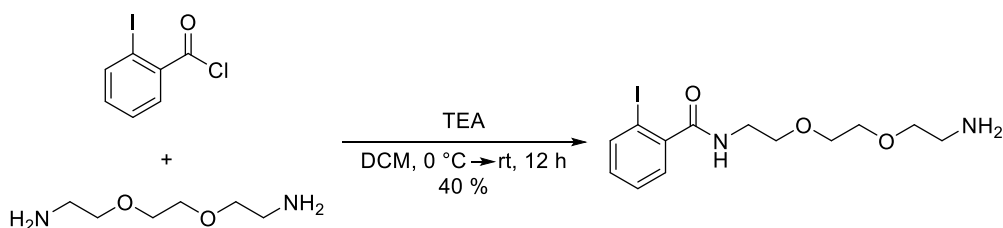
¹H NMR (400 MHz, DMSO-*d*₆) δ = 13.60 (s, 1H, COOH), 8.67 (d, *J* = 1.5 Hz, 1H, IC-CH_{ar}), 8.32 (dd, *J* = 7.8, 1.5 Hz, 1H, CH_{ar}-C-CH_{ar}), 8.12 (d, *J* = 7.8 Hz, 1H, IC-C-CH_{ar}).

¹³C NMR (101 MHz, DMSO-*d*₆) δ = 166.7 (IC-C-COOH), 166.0 (CH_{ar}-C-COOH), 147.4 (IC), 135.1 (C-COOH), 134.9 (C-COOH), 133.6 (CH_{ar}-C-CH_{ar}), 130.4 (IC-C-CH_{ar}), 125.8 (IC-CH_{ar}).

HRMS-ESI (*m/z*): berechnet für C₈H₄IO₆⁻ [M-H]⁻ : 322.9058; gefunden: 322.9059

IR (ATR, neat): ν = 2889 (w), 2526 (m), 1837 (m), 1694 (m), 1643 (w), 1549 (w), 1488 (m), 1427 (w), 1382 (w), 1317 (m), 1298 (m), 1249 (m), 1213 (w), 1125 (w), 953 (m), 921 (m), 868 (m), 807 (s), 775 (s), 754 (s), 722 (s), 698 (m), 660 (s).

N-(2-(2-(2-Aminoethoxy)ethoxy)ethyl)-2-iodbenzamid (**25**)



2,2'-(Ethylenedioxy)diethylamin (**24**) (2.96 mL, 20.2 mmol, 2.5 Äq) und Triethylamin (5.61 mL, 40.5 mmol, 5.0 Äq) wurden in DCM (100 mL) gelöst. Unter Eiskühlung wurde 2-Iodobenzoylchlorid (**23**) (gelöst in 40 mL DCM, 2.70 g, 10.1 mmol, 1 Äq) über 30 min zur Lösung getropft. Die Lösung wurde 2 h bei 0 °C und für 12 h bei Raumtemperatur gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und einer säulenchromatographischen Aufreinigung unterzogen („dry load“, Gradient, DCM/MeOH, 100:0 → 47:3). Das Produkt (1.51 g, 3.99 mmol, 40 %) wurde als leicht gelbliches Öl erhalten.

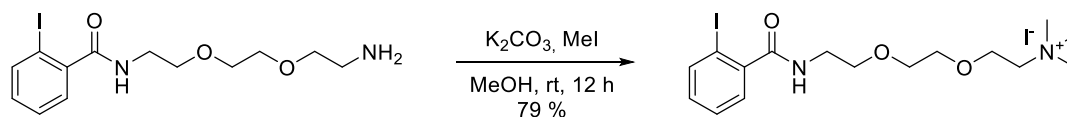
¹H NMR (400 MHz, MeOD) δ = 7.95 (dd, *J* = 8.0, 0.8 Hz, 1H, IC-CH_{ar}), 7.49 (td, *J* = 7.5, 1.1 Hz, 1H, IC-CH_{ar}-CH_{ar}), 7.41 (dd, *J* = 7.6, 1.6 Hz, 1H, IC-C-CH_{ar}), 7.26 – 7.16 (m, 1H, IC-C-CH_{ar}-CH_{ar}), 3.76 – 3.69 (m, 6H, CH₂), 3.59 (q, *J* = 5.3 Hz, 4H, CH₂), 2.84 (t, *J* = 5.3 Hz, 2H, NH₂-CH₂)

¹³C NMR (101 MHz, MeOD) δ = 172.6 (C=O), 144.1 (IC-C-CH_{ar}), 140.8 (IC-CH_{ar}), 132.0 (IC-C-CH_{ar}-CH_{ar}), 129.2 (IC-CH_{ar}-CH_{ar}), 129.0 (IC-C-CH_{ar}), 93.3 (C-I), 73.1 (CH₂), 71.4 (CH₂), 71.3 (CH₂), 70.3 (CH₂), 49.0 (CH₂), 41.9 (CH₂-NH₂), 40.8 (CH₂-NH)

HRMS-ESI (m/z): berechnet für C₁₃H₂₀IN₂O₃⁺ [M+H]⁺ : 379.0513; gefunden: 379.0512

IR (ATR, neat): ν = 3266 (w), 3051 (w), 2865 (w), 1641 (s), 1584 (m), 1535 (m), 1461 (m), 1424 (m), 1350 (m), 1303 (m), 1256 (w), 1098 (s), 1013 (s), 946 (m), 865 (m), 746 (s), 681 (s)

2-(2-(2-(2-Iodobenzamido)ethoxy)ethoxy)-N,N,N-trimethylethan-1-amoniumiodid (26)



25 (1.60 g, 4.23 mmol, 1.0 Äq) wurde in MeOH (40 mL) gelöst und K₂CO₃ (2.34 g, 16.92 mmol, 4.0 Äq) portionsweise hinzu gegeben. Zur Suspension wurde vorsichtig MeI (3.69 mL, 59.23 mmol, 14.0 Äq) getropft und über Nacht bei Raumtemperatur gerührt. Die Suspension wurde über einen Faltenfilter und Kieselgur gefiltert und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde einer säulenchromatographischen Aufreinigung unterzogen („dry load“, Gradient, DCM/MeOH, 93:7 → 9:1). Das Produkt (1.84 g, 3,36 mmol, 79 %) wurde als gelblicher Schaum erhalten.

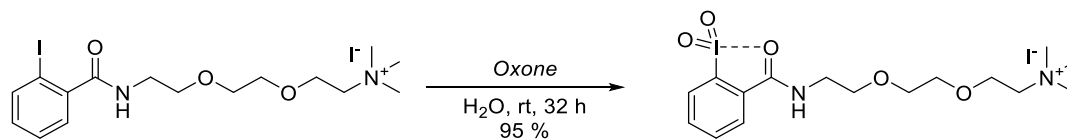
¹H NMR (400 MHz, D₂O) δ = 8.07 (d, J = 8.0 Hz, 1H, IC-CH_{ar}), 7.60 (t, J = 7.6 Hz, 1H, IC-CH_{ar}-CH_{ar}), 7.50 (dd, J = 7.6, 1.5 Hz, 1H, IC-C-CH_{ar}), 7.35 (td, J = 7.7, 1.7 Hz, 1H, IC-C-CH_{ar}-CH_{ar}), 4.06 (dq, J = 7.3, 2.8 Hz, 2H, CH₂-NMe₃), 3.86 (m, 6H, CH₂), 3.70 – 3.62 (m, 4H, CH₂), 3.25 (s, 9H, CH₃)

¹³C NMR (101 MHz, D₂O) δ = 172.7 (C=O), 141.0 (IC), 139.6 (IC-CH_{ar}), 131.7 (IC-CH_{ar}-CH_{ar}), 128.5 (IC-C-CH_{ar}-CH_{ar}), 127.7 (IC-C-CH_{ar}), 92.3 (IC-C-CH_{ar}), 69.8 (CH₂), 69.5 (CH₂), 68.7 (CH₂-CH₂-NH₂), 65.4 (CH₂-NMe₃), 64.4 (CH₂-CH₂-NMe₃), 54.2 (CH₃), 54.2 (CH₃), 54.1 (CH₃), 39.6 (CH₂-NH)

HRMS-ESI (m/z): berechnet für C₁₆H₂₆IN₂O₃⁺ [M+H]⁺ : 421.0983; gefunden: 421.0980

IR (ATR, neat): ν = 3384 (m), 2940 (w), 1636 (s), 1584 (m), 1537 (m), 1463 (m), 1423 (w), 1353 (w), 1307 (m), 1256 (w), 1094 (s), 1014(s), 952 (m), 873 (w), 823 (w), 751 (m), 723 (m), 680 (m)

**2-(2-(2-(2-Iodyl(V)benzamido)ethoxy)ethoxy)-N,N,N-trimethylethan-1-amoniumiodid
(IBX-Amid; 27)**



IBX Vorstufe **26** (200 mg, 365 μmol , 1.0 Äq) wurde in H_2O (12 mL) gelöst, portionsweise OXONE[®] (1.12 g, 3.65 mmol, 10.0 Äq) zu gegeben und für 32 h bei Raumtemperatur gerührt. Die Lösung wurde lyophilisiert und der Rückstand mit $\text{MeCN}/\text{H}_2\text{O}$ (2:1, 2 mL) zerrieben. Der zurückbleibende Feststoff wurde abgefiltert und das Filtrat erneut lyophilisiert. Der Rückstand wurde erneut mit $\text{MeCN}/\text{H}_2\text{O}$ (2:1, 1 mL) zerrieben, gefiltert und das Filtrat lyophilisiert. Das Produkt (200 mg, 345 μmol , 95 %) wurde als weißes Pulver erhalten.

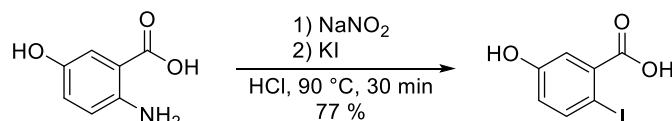
^1H NMR (400 MHz, D_2O) δ = 8.34 (dd, J = 8.0, 0.8 Hz, 1H, IC- CH_{ar}), 8.20 (dd, J = 7.8, 1.0 Hz, 1H, IC-C- CH_{ar}), 8.14 - 8.08 (m, 1H, IC- CH_{ar} - CH_{ar}), 7.92 (td, J = 7.7, 1.0 Hz, 1H, IC-C- CH_{ar} - C_{ar}), 3.92 (dq, J = 7.5, 2.8 Hz, 2H, $\text{CH}_2\text{-NMe}_3$), 3.81 (t, J = 5.1 Hz, 2H, CH_2), 3.76 - 3.68 (m, 6H, CH_2), 3.49 (m, 2H, $\text{CH}_2\text{-NMe}_3$), 3.10 (s, 9H, CH_3)

^{13}C NMR (101 MHz, D_2O) δ = 167.4 (C=O), 145.9 (IC), 134.9 (IC- CH_{ar} - CH_{ar}), 133.8 (IC-C- CH_{ar} - CH_{ar}), 128.0 (IC-C- CH_{ar}), 127.3 (IC-C- CH_{ar}), 123.3 (IC- CH_{ar}), 69.5 (CH_2), 69.4 (CH_2), 68.4 (CH_2), 65.2 ($\text{CH}_2\text{-NMe}_3$), 64.3 ($\text{CH}_2\text{-CH}_2\text{-NMe}_3$), 53.9 (CH_3), 53.9 (CH_3), 53.8 (CH_3), 40.3 ($\text{CH}_2\text{-NH}$)

HRMS-ESI (m/z): berechnet für $\text{C}_{16}\text{H}_{26}\text{IN}_2\text{O}_5^+$ $[\text{M}+\text{H}]^+$: 453.0881; gefunden: 453.0888

IR (ATR, neat): ν = 2875 (m), 2601 (w), 2360 (m), 1734 (w), 1717 (w), 1700 (w), 1684 (w), 1652 (w), 1623 (w), 1558 (w), 1279 (m), 1226 (m), 1153 (s), 1066 (s), 1000 (s), 871 (s), 846 (s), 738 (m)

5-Hydroxy-2-iodbenzoesäure (29)



5-Hydroxy-2-Aminobenzoessäure (**28**) (10.0 g, 65.3 mmol, 1.0 Äq) wurde in konzentrierter HCl (50 mL) gelöst und unter Eiskühlung NaNO_2 (gelöst in 20 mL H_2O , 5.86 g, 84.9 mmol, 1.3 Äq) innerhalb 30 min zu getropft. Die Reaktionsmischung wurde für 30 min bei 0 °C weitergerührt. Es

wurde KI (gelöst in 25 mL H₂O, 16.3 g, 98.0 mmol, 1.5 Äq) innerhalb 30 min zu getropft und für 30 min bei 90 °C erhitzt. Die Reaktionsmischung wurde mit H₂O (100 mL) verdünnt und mit EtOAc (3 x 400 mL) extrahiert. Die vereinigte organische Phase wurde mit einer Na₂S₂O₃ Lösung (10 % in H₂O, 3 x 200 mL) und einer NaCl Lösung (gesättigt, 300 mL) gewaschen und über MgSO₄ getrocknet. Das Lösungsmittel wurde *in vacuo* entfernt und das Produkt (13.2, 50.0 mmol, 77 %) als brauner Feststoff erhalten.

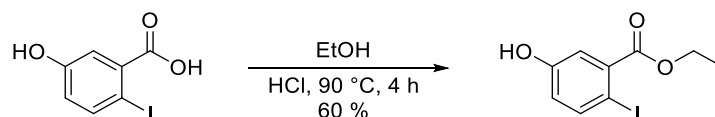
¹H NMR (400 MHz, MeOD): δ (ppm) = 7.70 (d, *J* = 8.6 Hz, 1H, IC-CH_{ar}), 7.19 (d, *J* = 3.0 Hz, 1H, COOH-C-CH_{ar}), 6.64 (dd, *J* = 8.6, 3.0 Hz, 1H, IC-CH_{ar}-CH_{ar})

¹³C NMR (101 MHz, MeOD): δ (ppm) = 170.1 (COOH), 158.9 (C-OH), 142.9 (IC-CH_{ar}), 138.5 (COOH-C), 121.2 (COOH-C-CH_{ar}), 118.8 (IC-CH_{ar}-CH_{ar}), 80.5 (IC)

HRMS-ESI (*m/z*): berechnet für C₇H₄IO₃⁻ [M-H]⁻: 262.9211, gefunden: 262.9213

IR (ATR, neat): ν = 3288 (w), 2810 (w), 2594 (w), 2355 (w), 2120 (w), 1916 (w), 1698 (m), 1584 (m), 1476 (w), 1426 (s), 1412 (m), 1304 (s), 1264 (s), 1240 (s), 1018 (m), 929 (m), 878 (m), 827 (s), 782 (m), 765 (s), 670 (m)

5-Hydroxy-2-iodbenzoesäureethylester (30)



5-Hydroxy-2-Iodbenzoesäure (**29**) (6.0 g, 22.7 mmol, 1.0 Äq), wurde in EtOH (80 mL, 1.36 mol, 60 Äq, abs.) gelöst und vorsichtig mit konzentrierter HCl (80 mL) versetzt. Die Lösung wurde für 4 h bei 90 °C erhitzt, danach mit H₂O (120 mL) verdünnt und mit EtOAc (3 x 300 mL) extrahiert. Die organische Phase wurde mit halbkonzentrierter NaCl Lösung (3 x 300 mL, bis pH ~ 7) gewaschen und über MgCO₃ getrocknet. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand einer säulenchromatographischen Aufreinigung unterzogen (Gradient, *i*Hex/EtOAc, 100:0 → 9:1). Das Produkt (3.94 g, 13.6 mmol, 60 %) wurde als farbloser Feststoff erhalten.

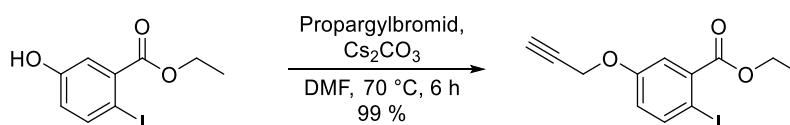
¹H NMR (400 MHz, CDCl₃) δ 7.78 (d, *J* = 8.6 Hz, 1H, IC-CH_{ar}), 7.33 (d, *J* = 3.0 Hz, 1H, COO-C-CH_{ar}), 6.71 (dd, *J* = 8.6, 3.1 Hz, 1H, IC-CH_{ar}-CH_{ar}), 4.39 (q, *J* = 7.1 Hz, 2H, CH₂), 1.40 (t, *J* = 7.1 Hz, 3H, CH₃)

¹³C NMR (101 MHz, CDCl₃) δ = 166.9 (COO), 156.0 (C-OH), 142.2 (IC-CH_{ar}), 136.3 (COO-C), 120.7 (IC-CH_{ar}-CH_{ar}), 118.5 (COO-C-CH_{ar}), 82.2 (IC), 62.2 (CH₂), 14.3 (CH₃)

HRMS-ESI (m/z): berechnet für $C_9H_8IO_3^-$ $[M-H]^-$: 290.9524, gefunden: 290.9527

IR (ATR, neat): ν = 3294 (m), 2976 (w), 2930 (w), 1880 (w), 1777 (w), 1692 (m), 1602 (w), 1566 (m), 1464 (m), 1441 (m), 1430 (m), 1396 (m), 1371 (m), 1322 (m), 1294 (s), 1258 (s), 1157 (m), 1144 (m), 1097 (s), 1009 (s), 952 (m), 943 (m), 890 (m), 868 (m), 817 (s), 780 (s), 756 (m), 675 (s)

5-*O*-(Prop-2-in-1-yloxy)-2-iodbenzoesäureethylester (**31**)



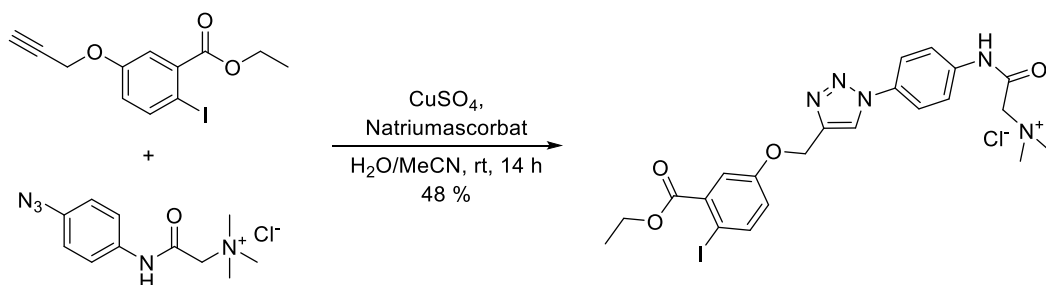
Es wurde **30** (2.40 g, 8.22 mmol, 1.0 Äq) in DMF (24 mL) gelöst und $CsCO_3$ (2.95 g, 9.04 mmol, 1.1 Äq) portionsweise zu gegeben. Die Mischung wurde für 10 min bei Raumtemperatur gerührt und Propargylbromid (80 % in Toluol, 1.09 mL, 11.5 mmol, 1.4 Äq) langsam zu getropft. Die Reaktionsmischung wurde für 4 h bei 70 °C erhitzt. Nach Abkühlung auf Raumtemperatur wurde die Mischung mit H_2O (100 mL) verdünnt und mit EtOAc (300 mL) extrahiert. Die organische Phase wurde mit einer halbgesättigten NaCl Lösung (3 x 150 mL) gewaschen und über $NaSO_4$ getrocknet. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand einer säulenchromatographischen Aufreinigung unterzogen (Gradient, *i*Hex/EtOAc, 100:0 \rightarrow 24:1). Das Produkt (2.68 g, 8.12 mmol, 99 %) wurde als leicht gelbliches Öl erhalten.

1H NMR (400 MHz, $CDCl_3$) δ = 7.86 (d, J = 8.7 Hz, 1H, IC- CH_{ar}), 7.43 (d, J = 3.1 Hz, 1H, COO-C- CH_{ar}), 6.83 (dd, J = 8.7, 3.1 Hz, 1H, IC- CH_{ar} - CH_{ar}), 4.70 (d, J = 2.4 Hz, 2H, CH_{Alkin} - CH_2), 4.40 (q, J = 7.1 Hz, 2H, CH_3 - CH_2), 2.54 (t, J = 2.4 Hz, 1H, CH_{Alkin}), 1.41 (t, J = 7.2 Hz, 3H, CH_3)

^{13}C NMR (101 MHz, $CDCl_3$) δ = 166.3 (COO), 157.5 (C-OH), 142.0 (IC- CH_{ar}), 136.4 (COO-C), 119.9 (IC- CH_{ar} - CH_{ar}), 117.9 (COO-C- CH_{ar}), 83.7 (IC), 77.8 (CH_{Alkin} -C- CH_2), 76.4 (CH_{Alkin}), 62.0 (CH_3 - CH_2), 56.2 (CH_{Alkin} -C- CH_2), 14.3 (CH_3)

HRMS-EI (m/z): berechnet für $C_{12}H_{11}IO_3$ $[M]$: 329.9753, gefunden: 329.9743

IR (ATR, neat): ν = 3435 (w), 3290 (w), 2979 (w), 2936 (w), 2872 (w), 2121 (w), 1721 (s), 1588 (w), 1564 (m), 1464 (m), 1406 (w), 1388 (w), 1366 (w), 1281 (s), 1244 (s), 1205 (s), 1172 (m), 1140 (m), 1098 (s), 1036 (s), 1008 (s), 935 (w), 905 (w), 888 (w), 859 (w), 812 (m), 776 (m), 674 (m)

SF⁺-IBX Vorstufe (33)

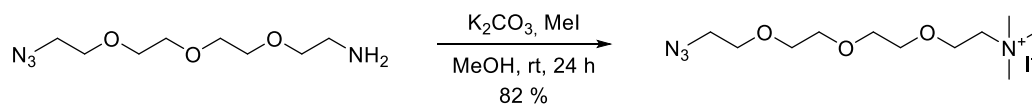
Das Alkin **31** (0.73 g, 2.22 mmol, 1.0 Äq) wurde in MeCN (18.5 mL) und der SF-Baustein **32** (in 22 mL H₂O gelöst, 0.60 g, 2.22 mmol, 1.0 Äq) zu getropft. Die Lösung wurde 2 Mal mit Argon entgast („Freeze-Pump-Thaw“), CuSO₄ (0.36 g, 2.22 mmol, 1.0 Äq) dazu gegeben und mit Argon ein weiteres Mal entgast. Die Cycloadditionsreaktion wurde durch Zugabe von Natriumascorbat (0.18 g, 0.89 mmol, 0.4 Äq) gestartet und erneut entgast. Die Reaktionsmischung wurde über Nacht bei Raumtemperatur gerührt und erneut Natriumascorbat (0.18 g, 0.89 mmol, 0.4 Äq) zu gegeben und 1 Mal entgast. Die Reaktion wurde für 10 h bei Raumtemperatur gerührt und das Lösungsmittel *in vacuo* entfernt. Der ölige Rückstand wurde einer säulenchromatographischen Aufreinigung unterzogen (isokratisch, DCM/MeOH, 9:1) um das Produkt (634 mg, 1.06 mmol, 48 %) in Form eines leicht grünlich weißen Feststoffs zu erhalten. Der Feststoff ist leider nur in DMSO löslich.

¹H NMR (400 MHz, DMSO-*d*₆): δ (ppm) = 8.92 (s, 1H, H_{Triazol}), 7.95 – 7.74 (m, 5H, H_{ar}), 7.39 (d, *J* = 3.0 Hz, 1H, H_{ar}), 7.06 (dd, *J* = 8.7, 3.1 Hz, 1H, H_{ar}), 5.29 (s, 2H, CH₂), 4.41 (s, 2H, CH₂), 4.31 (q, *J* = 7.1 Hz, 2H, CH₃-CH₂), 3.35 (s, 9H, N-CH₃), 1.32 (t, *J* = 7.1 Hz, 3H, CH₃)

¹³C NMR (101 MHz, DMSO-*d*₆): δ (ppm) = 166.4, 162.5, 157.8, 157.0, 143.4, 141.2, 138.1, 137.3, 123.0, 121.0, 120.6, 120.0, 116.8, 83.0, 79.0, 63.9, 61.5, 53.4, 14.0

HRMS-EI (*m/z*): berechnet für C₂₃H₂₇IN₅O₄⁺ [M]⁺: 564.1102, gefunden: 564.1103

IR (ATR, neat): ν = 3142 (w), 3056 (w), 2357 (w), 1691 (s), 1610 (w), 1557 (m), 1520 (m), 1473 (w), 1458 (w), 1412 (m), 1364 (w), 1286 (s), 1257 (s), 1230 (s), 1194 (m), 1151 (w), 1107 (m), 1056 (s), 1036 (m), 1009 (m), 990 (w), 920 (m), 835 (s), 808 (s), 781 (m), 756 (m), 678 (w)

11-Azid-3,6,9-trioxaundecan-1-*N,N,N*-trimethylamoniumiodid (36)

11-Azid-3,6,9-trioxaundecan-1-amin (**35**) (0.91 mL, 4.58 mmol, 1.0 Äq) wurde in MeOH (25 mL) gelöst und K_2CO_3 (2.41 g, 17.4 mmol, 3.8 Äq) portionsweise zu gegeben. Die Mischung wurde für 10 min bei Raumtemperatur gerührt und MeI (1.00 mL, 16.0 mmol, 3.5 Äq) langsam zu getropft. Die Reaktionsmischung wurde 12 h bei Raumtemperatur gerührt und erneut mit MeI (0.29 mL, 4.58 mmol, 1.0 Äq) versetzt. Die Reaktionsmischung wurde wieder für 12 h bei Raumtemperatur gerührt, über Kieselgur gefiltert und das Lösungsmittel *in vacuo* entfernt. Das Produkt (50 Gew. % in K_2CO_3 , 2.9 g, 3.73 mmol, 82 %) wurde als Mischung in K_2CO_3 erhalten. Das Produkt wurde durch lösen in DCM (25 mL) und filtern auch als Reinsubstanz (leicht gelbliches Öl) erhalten.

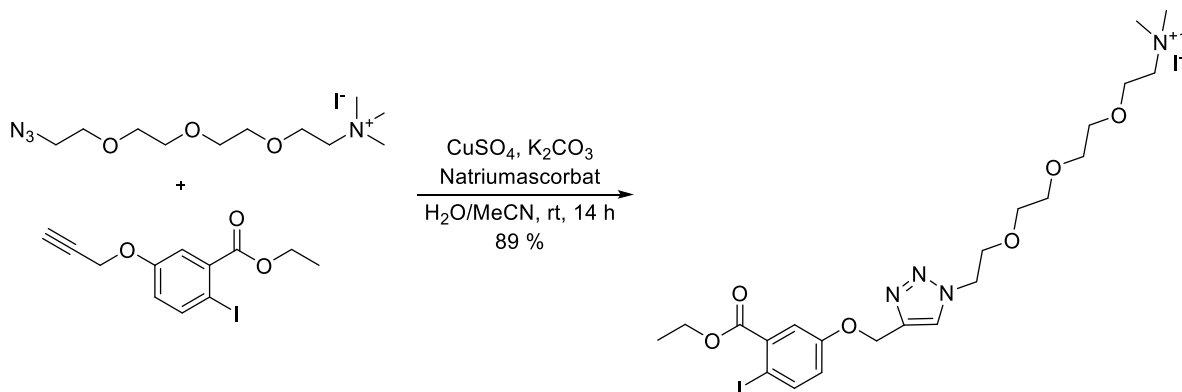
^1H NMR (400 MHz, CDCl_3) δ = 4.02 – 3.96 (m, 2H, CH_2), 3.93 (m, 2H, CH_2), 3.72 – 3.60 (m, 10H, CH_2), 3.49 (s, 9H, N- CH_3), 3.42 – 3.38 (m, 2H, CH_2)

^{13}C NMR (101 MHz, CDCl_3) δ = 70.8 (CH_2), 70.6 (CH_2), 70.6 (CH_2), 70.4 (CH_2), 70.1 (CH_2), 66.0 ($\text{CH}_2\text{-N-CH}_3$), 65.3 (CH_2), 55.1 (N- CH_3), 55.1 (N- CH_3), 55.1 (N- CH_3), 50.8 (CH_2)

HRMS-EI (m/z): berechnet für $\text{C}_{11}\text{H}_{25}\text{N}_4\text{O}_3^+$ [M] $^+$: 261.1921, gefunden: 261.1921

IR (ATR, neat): ν = 3424 (w), 3007 (w), 2868 (w), 2099 (m), 1623 (w), 1472 (w), 1444 (w), 1395 (w), 1353 (m), 1284 (m), 1247 (m), 1115 (s), 1094 (s), 1072 (s), 1007 (m), 969 (m), 947 (s), 920 (m), 875 (m), 863 (m), 828 (m), 701 (w), 662 (w)

11-N-(4-(O-(5-Oxy-2-iodbenzoesäureethylester)methyl)-1H-1,2,3-triazol-4-yl)-3,6,9-trioxaundecan-1-N,N,N-trimethylamoniumiodid (37)



Das Alkin **31** (0.50 g, 1.51 mmol, 1.0 eq) wurde in MeCN (14 mL) gelöst und das PEG⁺-Azid (**36**) (gelöst in 12 mL H₂O, 50 % K₂CO₃, 1.18 g, 1.51 mmol, 1.0 Äq) zur Lösung getropft. Die Lösung wurde mit Argon entgast (2x, „Freeze-Pump-Thaw“), CuSO₄ (0.21 g, 1.51 mmol, 1.0 Äq) dazugegeben und mit Argon ein weiteres Mal entgast. Die Cycloadditionsreaktion wurde durch Zugabe von Natriumascorbat (0.12 g, 0.61 mmol, 0.4 Äq) gestartet und die Lösung erneut entgast. Die Reaktionsmischung wurde über Nacht bei Raumtemperatur gerührt und das Lösungsmittel *in vacuo* entfernt. Der ölige Rückstand wurde einer säulenchromatographischen Aufreinigung unterzogen (isokratisch, DCM/MeOH, 9:1) um das Produkt (0.96 g, 1.34 mmol, 89 %) in Form eines orange-braunen Öls zu erhalten.

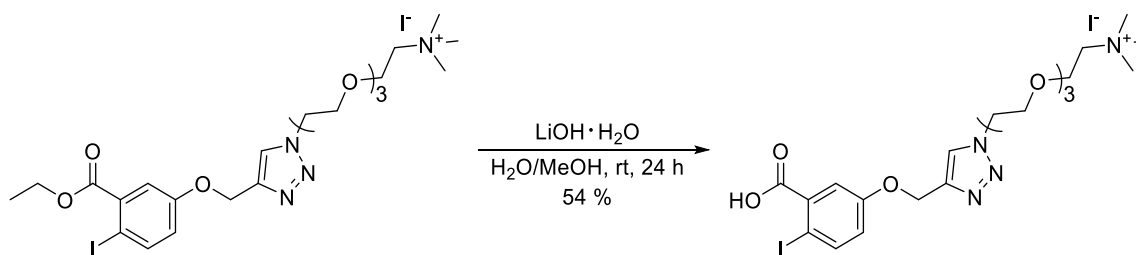
¹H NMR (400 MHz, D₂O) δ = 8.19 (s, 1H, **H**_{Triazol}), 7.86 (d, *J* = 8.6 Hz, 1H, IC-CH_{ar}), 7.38 (d, *J* = 2.8 Hz, 1H, COO-C-CH_{ar}), 6.96 (dd, *J* = 8.6, 3.2 Hz, 1H, IC-CH_{ar}-CH_{ar}), 5.33 (s, 2H, O-CH₂-C_{Triazol}), 4.71 – 4.66 (t, *J* = 4.6 Hz, 2H, CH₂), 4.41 (q, *J* = 7.0 Hz, 2H, CH₃-CH₂-COO), 3.98 (t, *J* = 4.6 Hz, 2H, CH₂), 3.93 – 3.87 (m, 2H, N-CH₂-CH₂), 3.62 – 3.46 (m, 10H, CH₂), 3.17 (s, 9H, N-CH₃), 1.41 (t, *J* = 7.1 Hz, 3H, CH₃-CH₂-COO)

¹³C NMR (101 MHz, D₂O) δ = 168.3 (COO), 157.3 (CH₂O-C-CH_{ar}), 143.0 (schwach, C_{Triazol}), 141.8 (IC-CH_{ar}), 136.3 (COO-C-CH_{ar}), 125.6 (CH_{Triazol}), 120.3 (IC-CH_{ar}-CH_{ar}), 117.4 (COO-C-CH_{ar}), 82.6 (IC), 69.7 (CH₂), 69.5 (CH₂), 69.4 (CH₂), 69.4 (CH₂), 68.7 (CH₂), 65.2 (t, CH₃-N-CH₂), 64.2 (CH₂), 63.1 (CH₃-CH₂-COO), 61.2 (O-CH₂-C_{Triazol}), 53.9 (CH₃-N), 53.8 (CH₃-N), 53.8 (CH₃-N), 50.2 (CH₂), 13.4 (CH₃-CH₂-COO).

HRMS-EI (*m/z*): berechnet für C₂₃H₃₆IN₄O₆⁺ [M]⁺: 591.1674, gefunden: 591.1676.

IR (ATR, neat): ν = 3399 (w), 2908 (w), 1713 (m), 1660 (w), 1621 (w), 1586 (w), 1563 (w), 1464 (m), 1405 (w), 1365 (w), 1283 (s), 1248 (m), 1215 (s), 1099 (s), 1063 (s), 1023 (s), 1006 (s), 951 (m), 928 (m), 873 (w), 820 (m), 778 (m), 759 (m)

11-(4-(O-(5-Oxy-2-iodbenzoesäure)methyl)-1*H*-1,2,3-triazol-4-yl)-3,6,9-trioxaundecan-1-*N,N,N*-trimethylamoniumiodid (38)



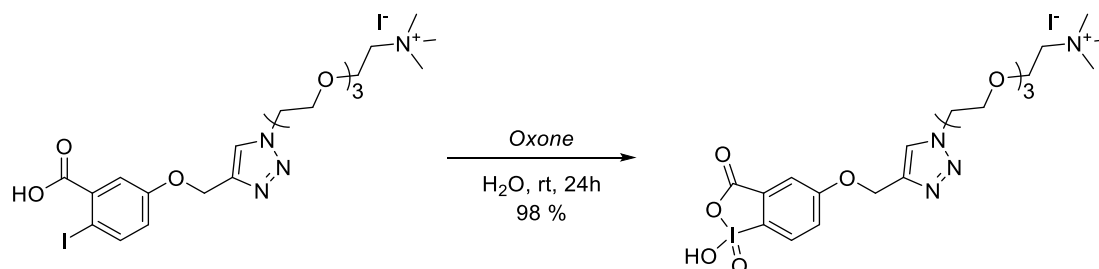
37 (0.56 g, 0.78 mmol, 1.0 Äq) wurde in MeOH/H₂O (3:1, 55 mL) gelöst und Lithiumhydroxid Monohydrat (0.20 g, 4.70 mmol, 6.0 Äq) zur Lösung gegeben. Die Mischung wurde für 24 h bei Raumtemperatur gerührt. Die Suspension wurde über Kieselgur gefiltert und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde erneut in H₂O (15 mL) gelöst und in 5 Läufen über semipräparative HPLC (Gradient über 38 min; Puffer A = H₂O; Puffer B = 80 % MeCN_{aq}; 0 →_{13min} 0 →_{25min} 30 % Puffer B) einer Aufreinigung unterzogen. Das Produkt (0.20 g, 0.28 mmol, 54 %) wurde als farbloses Öl erhalten.

¹H NMR (599 MHz, D₂O) δ = 8.16 (s, 1H, **H**_{Triazol}), 7.77 (d, *J* = 8.7 Hz, 1H, IC-**CH**_{ar}), 7.04 (s, 1H, COO-C-**CH**_{ar}), 6.81 (d, *J* = 8.7 Hz, 1H, IC-**CH**_{ar}-**CH**_{ar}), 5.27 (s, 2H, O-**CH**₂-C_{Triazol}), 4.64 (t, *J* = 4.8 Hz, 2H, **CH**₂), 3.97 (t, *J* = 4.8 Hz, 2H, **CH**₂), 3.89 (s, 2H, **CH**₂), 3.61 – 3.52 (m, 10H, **CH**₂), 3.16 (s, 9H, N-**CH**₃)

¹³C NMR (101 MHz, D₂O) δ = 176.5 (COO), 157.4 (CH₂-O-C-**CH**_{ar}), 147.4 (C), 142.9 (C), 140.1 (IC-**CH**_{ar}), 125.5 (CH_{Triazol}), 116.5 (IC-**CH**_{ar}-**CH**_{ar}), 113.6 (COO-C-**CH**_{ar}), 80.4 (IC), 69.7 (CH₂), 69.4 (CH₂), 69.3 (CH₂), 69.3 (CH₂), 68.6 (CH₂), 65.1 (t, CH₃-N-CH₂), 64.2 (CH₂), 61.1 (O-CH₂-C_{Triazol}), 53.8 (CH₃-N), 53.8 (CH₃-N), 53.7 (CH₃-N), 50.1 (CH₂)

HRMS-EI (*m/z*): berechnet für C₂₁H₃₂IN₄O₆⁺ [M]⁺: 563.1361, gefunden: 563.1361

11-(4-(O-(5-Oxy-2-iod(V)benzoesäure)methyl)-1H-1,2,3-triazol-4-yl)-3,6,9-trioxaundecan-1-N,N,N-trimethylamoniumiodid (PEG⁺-IBX; 39)



38 (146 mg, 211 μ mol, 1.0 Äq) wurde in H₂O (8.7 mL) gelöst, Oxon[®] (648 mg, 2.11 mmol, 10.0 Äq) in Portionen zur Lösung gegeben und die Reaktionslösung 12 h bei Raumtemperatur gerührt. Das Lösungsmittel wurde *in vacuo* entfernt und der Überstand mit MeCN/H₂O (2:1, 30 mL) zerrieben. Die Mischung wurde gefiltert und das Lösungsmittel des Filtrats *in vacuo* entfernt. Der Überstand wurde ein weiteres Mal mit MeCN/H₂O (2:1, 15 mL) zerrieben, gefiltert und das Lösungsmittel des Filtrates *in vacuo* entfernt. Das Produkt (150 mg, 208 μ mol, 98 %) wurde als leicht gelbliches Öl erhalten. Die Stabilität ist ungelöst bei Raumtemperatur begrenzt.

¹H NMR (400 MHz, D₂O) δ = 8.15 (s, 1H, CH_{Triazol}), 8.08 (d, J = 8.9 Hz, 1H, IC-CH_{ar}), 7.64 (d, J = 2.7 Hz, 1H, COO-C-CH_{ar}), 7.55 (dd, J = 8.9, 2.7 Hz, 1H, IC-CH_{ar}-CH_{ar}), 5.35 (s, 2H, O-CH₂-C_{Triazol}), 4.59 – 4.52 (m, 2H, CH₂), 3.92 – 3.86 (m, 2H, CH₂), 3.80 (dq, J = 7.5, 2.8 Hz, 2H, CH₂), 3.53 – 3.42 (m, 10H, CH₂), 3.06 (s, 9H, N-CH₃)

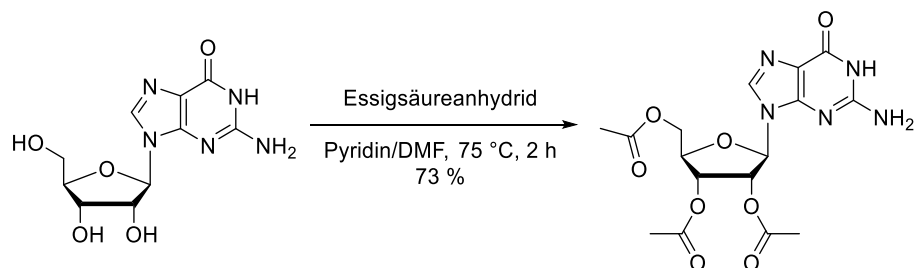
¹³C NMR (101 MHz, D₂O) δ = 170.3 (COO), 162.3 (CH₂O-C-CH_{ar}), 142.5 (C_{Triazol}), 136.8 (IC), 133.0 (COO-C-CH_{ar}), 126.1 (IC-CH_{ar}), 125.8 (CH_{Triazol}), 120.4 (IC-CH_{ar}-CH_{ar}), 116.8 (COO-C-CH_{ar}), 69.7 (CH₂), 69.5 (CH₂), 69.4 (CH₂), 69.3 (CH₂), 68.7 (CH₂), 65.2 (CH₃-N-CH₂), 64.3, 61.7 (O-CH₂-C_{Triazol}), 53.9 (N-CH₃), 53.9 (N-CH₃), 53.8 (N-CH₃), 50.2 (CH₂)

HRMS-EI (m/z): berechnet für C₂₁H₃₂IN₄O₈⁺ [M]⁺: 595.1259, gefunden: 595.1271

IR (ATR, neat): ν = 2880 (w), 2482 (w), 2128 (w), 1730 (m), 1572 (m), 1475 (w), 1151 (s), 1050 (s), 1018 (s), 870 (s), 847 (s)

8.1.2.3. Synthese von d3-ms²A

2',3',5'-O-Acetylguanosin (41)



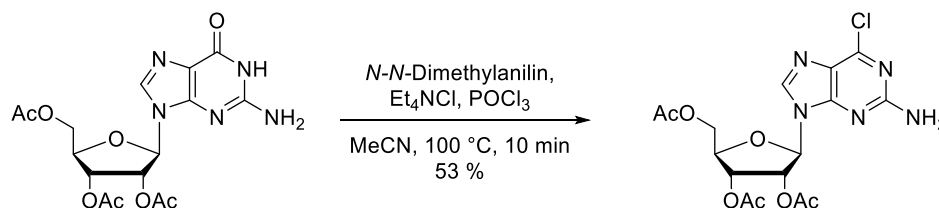
Guanosin **40** (5.0 g, 17.7 mmol, 1.0 Äq) wurde in einer Mischung aus Pyridin (10 mL) und DMF (12 mL) gelöst und über 10 min Essigsäureanhydrid (20.0 mL, 211.8 mmol, 12 Äq) hinzugegeben. Die Reaktionsmischung für 2 h bei 75 °C gerührt und mit Ethanol (1 mL) versetzt. Das Lösungsmittel wurde *in vacuo* entfernt und der Überstand in *iso*-Propanol einer Umkristallisation unterzogen. Das Produkt (5.3 g, 13.0 mmol, 73 %) wurde als leicht gelbliches Pulver erhalten.

¹H NMR (400 MHz, DMSO-*d*₆) δ = 10.74 (s, 1H, **NH**_{1,Base}), 7.93 (s, 1H, **H**_{8,Base}), 6.53 (s, 2H, **NH**₂), 5.99 (d, *J* = 6.1 Hz, 1H, **CH**_{1'}), 5.79 (t, *J* = 6.0 Hz, 1H, **CH**_{2'}), 5.49 (dd, *J* = 5.9, 4.1 Hz, 1H, **CH**_{3'}), 4.43 – 4.20 (m, 3H, , **CH**_{4'+5'}), 2.11 (s, 3H, **CH**₃, bei 3'), 2.04 (s, 3H, **CH**₃, bei 5'), 2.04 (s, 3H, **CH**₃, bei 2').

¹³C NMR (101 MHz, DMSO-*d*₆) δ = 170.1 (**COO** bei 5'), 169.4 (**COO** bei 3'), 169.3 (**COO** bei 2'), 156.6 (**C**_{6,Base}), 153.9 (**C**_{2,Base}), 151.1 (**C**_{4,Base}), 135.6 (**CH**_{8,Base}), 116.8 (**C**_{9,Base}), 84.4 (**CH**_{1'}), 79.5 (**CH**_{4'}), 72.0 (**CH**_{2'}), 70.3 (**CH**_{3'}), 63.1 (**CH**_{5'}), 20.5 (**CH**₃), 20.4 (**CH**₃), 20.2 (**CH**₃).

HRMS-ESI (*m/z*): berechnet für C₁₆H₂₀N₅O₈⁺ [**M+H**]⁺: 410.1306; gefunden: 410.1306

IR (ATR, neat): ν = 3424 (w), 3306 (w), 3153 (w), 2719 (w), 1745 (m), 1725 (w), 1698 (s), 1645 (m), 1594 (m), 1542 (w), 1525 (w), 1481 (w), 1431 (w), 1379 (m), 1366 (m), 1334 (w), 1229 (s), 1174 (m), 1141 (w), 1094 (m), 1040 (s), 1016 (m), 964 (w), 920 (m), 903 (m), 809 (m), 784 (m), 743 (w), 686 (m)

6-Chlor-2',3',5'-O-acetylguanosin (42)

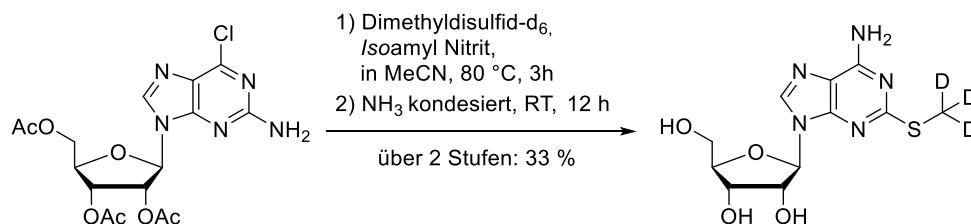
Es wurde Acetylgeschütztes Guanosin (2.0 g, 4.9 mmol, 1.0 Äq) mit Tetraethylammonium-chlorid (1.6 g, 9.8 mmol, 2.0 Äq) und *N,N*-Dimethylanilin (617 µL, 4.9 mmol, 1.0 Äq) in MeCN (10 mL) gelöst. Zur Lösung wurde unter Eiskühlung innerhalb von 10 min POCl₃ (2.7 mL, 29.3 mmol, 6.0 Äq) getropft. Die Lösung wurde 10 min bei 100 °C gerührt und anschließend das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde in ddH₂O (100 mL) aufgenommen und 5 Mal mit Chloroform (je 100 mL) extrahiert. Die vereinigte Organische Phase wurde mit einer gesättigten NaHCO₃ (aq., 100 mL) gewaschen und über Na₂SO₄ getrocknet. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand einer säulenchromatographischen Aufreinigung unterzogen (Gradient, DCM/MeOH, 50:1 → 97:3). Das Produkt (1.1 g, 2.6 mmol, 53 %) wurde als gelber Schaum erhalten.

¹H NMR (400 MHz, DMSO-*d*₆) δ = 8.36 (s, 1H, **H**_{8,Base}), 7.05 (s, 2H, **NH**₂), 6.10 (d, *J* = 5.9 Hz, 1H, **CH**_{1'}), 5.87 (t, *J* = 5.9 Hz, 1H, **CH**_{2'}), 5.54 (dd, *J* = 5.9, 4.2 Hz, 1H, **CH**_{3'}), 4.43 – 4.25 (m, 3H, **CH**_{4'+5'}), 2.12 (s, 3H, **CH**₃, bei 3'), 2.03 (s, 3H, **CH**₃, bei 5'), 2.03 (s, 3H, **CH**₃, bei 2').

¹³C NMR (101 MHz, DMSO-*d*₆) δ = 170.0 (**COO** bei 5'), 169.4 (**COO** bei 3'), 169.2 (**COO** bei 2'), 159.9 (**C**_{6,Base}), 153.6 (**C**_{2,Base}), 149.9 (**C**_{4,Base}), 141.2 (**CH**_{8,Base}), 123.5, 84.9 (**CH**_{1'}), 79.7 (**CH**_{4'}), 71.9 (**CH**_{2'}), 70.2 (**CH**_{3'}), 62.9 (**CH**_{5'}), 20.5 (**CH**₃), 20.4 (**CH**₃), 20.2 (**CH**₃).

HRMS-ESI (*m/z*): berechnet für C₁₆H₁₉ClN₅O₇⁺ [**M+H**]⁺: 428.0968; gefunden: 428.0971

IR (ATR, neat): ν = 3478 (w), 3332 (w), 3203 (w), 1741 (s), 1611 (m), 1561 (m), 1512 (w), 1469 (m), 1431 (w), 1406 (m), 1372 (m), 1322 (w), 1213 (s), 1146 (m), 1090 (m), 1043 (s), 999 (m), 953 (m), 906 (s), 865 (m), 802 (w), 786 (m), 732 (w)

d3-ms²A (43)

Das chlorierte Nukleosid **42** (171 mg, 0.4 mmol, 2.0 Äq) wurde in MeCN (800 µL) gelöst und Dimethyldisulfid-d₆ (18 µL, 0.2 mmol, 1.0 Äq) zu getropft. Die Lösung wurde mit Isoamyl Nitrit (107 µL, 0.8 mmol, 4.0 Äq) versetzt und 5 min bei 70 °C gerührt. Die Reaktionsmischung wurde in ein Druckrohr überführt und das Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde mit NH₃ (2 mL) gelöst und die Lösung über Nacht bei Raumtemperatur gerührt. Die Reaktionsmischung wurde getrocknet und der bräunliche Rückstand einer säulenchromatographischen Aufreinigung unterzogen (Gradient, DCM/MeOH, 19:1 → 9:1). Das Produkt wurde erneut in MeCN (300 µL) aufgenommen, mit ddH₂O (2.7 mL) verdünnt und über semipräparative HPLC (Gradient über 45 min; Puffer A = H₂O; Puffer B = 80 % MeCN_{aq}; 10 →_{45min} 25 % Puffer B) einer erneuten Aufreinigung unterzogen. Das Produkt (21 mg, 66.4 µmol, 33 %) wurde als weißes Pulver erhalten.

¹H NMR (400 MHz, DMSO-*d*₆) δ = 8.23 (s, 1H, H_{8,Base}), 7.38 (s, 2H, NH₂), 5.83 (d, *J* = 6.0 Hz, 1H), 5.44 (d, *J* = 6.0 Hz, 1H), 5.20 (d, *J* = 4.8 Hz, 1H), 5.02 (t, *J* = 5.6 Hz, 1H), 4.61 (q, *J* = 5.8 Hz, 1H), 4.14 (q, *J* = 4.7 Hz, 1H), 3.91 (q, *J* = 4.2 Hz, 1H), 3.67 – 3.49 (m, 2H, H₅).

¹³C NMR (101 MHz, DMSO-*d*₆, schwache Signale) δ 156.8, 155.5, 150.2, 138.8, 116.9, 87.2, 85.4, 73.2, 70.5, 61.6, 52.7, 22.0.

HRMS-ESI (*m/z*): berechnet für C₁₁H₁₃D₃N₅O₄S⁺ [M+H]⁺: 317.1106; gefunden: 317.1105

8.2. Oligonukleotidsynthese

8.2.1. Festphasensynthese

Die Festphasensynthese von DNA Oligonukleotiden, RNA Oligonukleotiden und Hybrid Oligonukleotiden (RNA und DNA) erfolgte auf einem *DNA/RNA Synthesizer Model 394* von der Firma *Applied Biosystems*. RNA Synthesen wurden im 1 µmol Maßstab und DNA und Hybrid Synthesen im 200 nmol Maßstab durchgeführt. Für die RNA Synthesen wurden 2'-TBDMS-geschützte Standard Phosphoramidite (Bz-rA, Ac-rC, dmf-rG, rU) und CPG-Träger (*SynBase*; Porengröße 1000 Å; Partikelgröße 110 µm) von der Firma *LinkTech* genutzt. Für die DNA und Hybridsynthese wurden Standard Phosphoramidite (Bz-dA, Ac-dC, iBu-dG und T von *LinkTech* und *Sigma-Aldrich*) und Low Volume Polystyrene-Träger (*GlenResearch*) sowie CPG-Träger (*SynBase*; Porengröße 1000 Å; Partikelgröße 110 µm) genutzt. Für die Synthese und das Anlösen der Phosphoramidite wurde Acetonitril (*Rotidry*, Wassergehalt < 10 ppm) von der Firma *Carl Roth* genutzt. Standard Phosphoramidite wurden mit der Konzentration 0.1 M und Sonder-Phosphoramidite mit der Konzentration 0.1 – 0.15 M gelöst. Als Aktivator wurde BTT (RNA, 0.3 M in Acetonitril, *LinkTech*) und Aktivator 42 (DNA und Hybride, 0.25 M in Acetonitril, 5-[3,5-Bis(trifluoromethyl)phenyl]-1*H*-tetrazol, *Sigma-Aldrich*) genutzt. Für das *deblocking* (*emp Biotech*, 3 % (v/v) Dichloressigsäure in DCM), die *oxidation* (*J.T. Baker*) und das *capping A/B* (*J.T. Baker*) wurden vorgemischte Lösungen genutzt. Die Kopplungszeiten der Standard Phosphoramidite betrugen 10 min (RNA) bzw. 45 s (DNA) und wurden für Sonder-Phosphoramidite auf 3-5 min (DNA) verlängert. Im letzten Syntheseschritt wurde grundsätzlich die letzte 5'-DMT-Schutzgruppe entfernt (*DMT-off*), dabei wurde für Biotin, zur vollständigen Entschützung, die Zeit für das *deblocking* auf 45 s (DNA & RNA) verlängert. Die Kartuschen wurden nach Abschluss der Synthese für 5 min am Hochvakuum getrocknet. (Lagerung ist nun bei 4 °C - -20 °C über mehrere Monate möglich)

8.2.2. Entschützung und Aufreinigung

Allgemein

Das trockene Trägermaterial aus der DNA/RNA Festphasensynthese wurde in ein 15 mL Reaktionsgefäß (*Falcon*) überführt. Es wurde eine Mischung aus NH₄OH (500 µL, 28 % in H₂O) und Methylamin (500 µL, 40 % in H₂O) frisch hergestellt (AMA) und zum Trägermaterial gegeben. Die Suspension wurde für 5 min (25 °C bei 600 rpm) und für 15 min (60 °C bei 600 rpm) inkubiert. Die AMA Entschützung wurde mit Eis auf 0 °C gekühlt und für 5 min zentrifugiert (1150 g, 4 °C), der Überstand abgenommen und das Trägermaterial mit ddH₂O gewaschen (1 mL und 0.5 mL, 1150 g, 4 °C). Die vereinigten Überstände wurden *in vacuo* getrocknet, mit DMSO (100 µL) aufgenommen und in einem 2 mL Reaktionsgefäß (*Eppendorf Tube*) mit Et₃N·3HF (125 µL, *Sigma-Aldrich*) versetzt. Die Reaktionsmischung wurde für 90 min bei 65 °C und 400 rpm inkubiert, abgekühlt und mit Natriumacetat (3.0 M, pH = 5.5, 25 µL) versetzt. Die Lösung wurde 1 min geschüttelt, *n*BuOH (1 mL) zugegeben und über Nacht bei -80 °C gekühlt. Das Präzipitat wurde zentrifugiert (1 h, 4 °C, 21130 g)

und der Überstand verworfen. Das Pellet wurde in HNEt₃OAc-Puffer (1 mL, 1 M, pH = 7) gelöst und mittels semipräparativer HPLC gereinigt.

AAF-dG (Ultra Ultra Mild): ^[241]

Die Synthese von AAF-dG haltigen Oligonukleotide wurde unter *ultramild* Synthese-Bedingungen der Firma *Ella Biotech* durchgeführt (das AAF-dG Phosphoramidite wurde im Rahmen dieser Doktorarbeit aus Vorstufen (*Dr. Nina Simon*) synthetisiert). Das Oligonukleotid wurde „DMT-on“ und immobilisiert am Trägermaterial (Q-Linker auf CPG von *Glen Research*, Porengröße 500 Å; Partikelgröße 110 µm) geliefert. Das Trägermaterial wurde in ein 2 mL Reaktionsgefäß (*eppendorf*) überführt und mit 1.4 mL *ultra-ultramild*-Puffer (0.25 M β-Mercaptoethanol, 10 % (v/v) Diisopropylamin in MeOH) für 18 h bei 55 °C und 600 rpm entschützt. Der Überstand wurde gefiltert und das Trägermaterial mit ddH₂O (2 x 200 µL) gewaschen. Die vereinigten Überstände wurden mittels *SpeedVac* eingengt, mit HNEt₃OAc-Puffer (1 mL, 1 M, pH = 7) versetzt und mittels semipräparativer HPLC gereinigt. Die gereinigten Fraktionen wurden vereint und *in vacuo* getrocknet. Das Pellet wurde in Essigsäure (1 mL, 80 % in ddH₂O) gelöst und für 40 min bei 25 °C und 600 rpm entschützt. Das Lösungsmittel wurde mittels *SpeedVac* um die Hälfte eingengt und mit HNEt₃OAc-Puffer (1 mL, 1 M, pH = 7) versetzt und mittels semipräparativer HPLC gereinigt.

hmdC oder cadC:

Das Trägermaterial wurde mit 1 mL einer 0.4 M NaOH Lösung in MeOH/ddH₂O (4:1) versetzt und 17 h bei 25 °C und 600 rpm inkubiert. Der Überstand wurde gefiltert und das Trägermaterial mit ddH₂O (2 x 450 µL) gewaschen. Der vereinigte Überstand wurde mittels *SpeedVac* konzentriert und mit Natriumacetat (3 M, 1/10 Volumen) und EtOH (4 x Volumen) über Nacht bei -20 °C gefällt. Das Präzipitat wurde für 1 h, bei 4 ° und 21130 g pelletiert, mit HNEt₃OAc-Puffer (1 mL, 1 M, pH = 7) versetzt und mittels semipräparativer HPLC gereinigt.

fdC:

Das Trägermaterial wurde mit 1 mL einer 0.4 M NaOH Lösung in MeOH/ddH₂O (4:1) versetzt und 17 h bei 25 °C und 600 rpm inkubiert. Der Überstand wurde gefiltert und das Trägermaterial mit ddH₂O (2 x 450 µL) gewaschen. Der vereinigte Überstand wurde mittels *SpeedVac* konzentriert und mit Natriumacetat (3 M, 1/10 Volumen) und EtOH (4 x Volumen) über Nacht bei -20 °C gefällt. Das Präzipitat wurde für 1 h, bei 4 ° und 21130 g pelletiert, zur Acetylentschätzung mit 1 mL Essigsäure (80 % in ddH₂O) versetzt und für 6 h bei 20 °C inkubiert. Die Lösung wurde mit HNEt₃OAc-Puffer (1 mL, 1 M, pH = 7) versetzt und mittels semipräparativer HPLC gereinigt.

mdC, hmdC, fdC und cadC (Hairpins):

Das Trägermaterial wurde mit 1 mL einer 0.4 M NaOH Lösung in MeOH/ddH₂O (4:1) versetzt und 26 h bei 25 °C und 700 rpm inkubiert. Die Suspension wurde zentrifugiert (4 min, 25 °C, 21130 g), der Überstand abgenommen und der Träger mit ddH₂O (1 mL und 500 µL) gewaschen. Der vereinigte

Überstand wurde mittels *SpeedVac* eingeeengt (500 µL Restvolumen) und erneut gefiltert. Die Lösung wurde mit Natriumacetat (3 M, 1/10 Volumen) und EtOH (4 x Volumen) über Nacht bei -20 °C gefällt. Das Präzipitat wurde pelletiert (1 h, 4 °C, 21130 g), mittels Sep-Pak Kartuschen (C18 Classic Plus, *Waters*) entsalzt und anschließend lyophilisiert. Das Pellet wurde zur Acetylentzschützung in 100 µL Essigsäure (80 % in ddH₂O) gelöst und für 6 h bei 20 °C und 700 rpm inkubiert. Das Oligonukleotid wurde erneut mit Natriumacetat (3 M, 1/10 Volumen) und EtOH (4 x Volumen) über Nacht bei -20 °C gefällt. Das Präzipitat wurde pelletiert (1 h, 4 °C, 21130 g), mit HNEt₃OAc-Puffer (1 mL, 1 M, pH = 7) und ddH₂O (2 x 1 mL) angelöst und mittels semipräparativer HPLC gereinigt. Die semipräparative HPLC wurde ein weiteres Mal wiederholt, um die Reinheit zu erhöhen.

Entsalzung von Oligonukleotiden

Alle synthetisierten Oligonukleotide wurden nach der finalen HPLC Aufreinigung mittels Sep-Pak Kartuschen (C18 Classic Plus, *Waters*) entsalzt. Dafür wurde die Sep-Pak Kartusche mit MeCN (10 mL) gewaschen und mit HNEt₃OAc-Puffer (10 mL, 0.1 M, pH = 7) äquilibriert. Das Oligonukleotid wurde in HNEt₃OAc-Puffer (2 mL, 0.1 M, pH = 7) gelöst und in 2 min auf die Kartusche aufgetragen. Die Kartusche wurde nun mit ddH₂O (10 mL) gewaschen, um den verbleibenden HNEt₃OAc-Puffer zu entfernen. Anschließend wurde das Oligonukleotid mit MeCN/ddH₂O (80:20, 6 mL) eluiert.

8.2.3. Charakterisierung

Die finale Identität und Reinheit der Oligonukleotide wurde durch analytischer HPLC und MALDI-ToF-MS (*autoflex II Unit* von *Bruker*) Messungen überprüft. Die Oligonukleotide wurden dabei mit HPA-Matrix (Tabelle 12) co-kristallisiert. Die Konzentrationsbestimmung wurde mittels NanoDrop ND-1000 UV/Vis-Spectrophotometer (*PreQlab*) durchgeführt.

8.2.4. Oligonukleotid Sequenzen

Tabelle 12: Sequenzen aller in dieser Arbeit synthetisierten Oligodesoxyribonukleotide. Modifizierte Basen wurden Fett markiert. B = 5'-Biotin (LinkTech/ Stöckl); W = A/T; D = A/G/T; N = A/C/G/T; F = FAM6.

ODN #	Sequenz
ODN 1	5'-B N fd CN fd CNC-3'
ODN 2	5'-B NCN CNC-3'
ODN 3	5'-B NNN fd CNN NA-3'
ODN 4	5'-B NNN CNN NA-3'
ODN 7	5'-B TTC ATC ATC GTC GTC GTC GTC ATC ATC TT-3'
ODN 8'	5'-AAG ATG ATG A hmd CG A hmd CG A hmd CG A hmd CG ATG ATG AA-3'
ODN 9	5'-B TTC ATC A Thmd C G Thmd C G Thmd C G Thmd C GTC ATC ATC TT-3'
ODN 10	5'-B TTC ATC AT fd C GT fd C GT fd C GT fd C GTC ATC ATC TT-3'
ODN 11	5'-B TTC ATC AT cad C GT cad C GT cad C GT cad C GTC ATC ATC TT-3'

ODN 12	5'–CGC GmdC GhmdC ATD DDD DDD DAT CG fdC GcadCC G–3'
ODN 13	5'–TNA GCG mdCGhmdC GAT DDD DDD DDA TCG fdCGcadC GC–3'
ODN 14	5'–CCG GCG GmdCG hmdCGA TDD DDD DDD ATC GfdCG cadCCG –3'
ODN 15	5'–GWC GGG mdCGhmdC GAT DDD DDD DDA TCG fdCGcadC CC–3'
ODN 16	5'–GGG mdCGhmdC GAT DDD DDD DDA TCG fdCGcadC CCT–3'
ODN 17	5'–GTA ATG hmdCGC TAG G–3'
ODN 18	5'–CCA GGG GAG TGT CGA TAAF-dGC ACA GAT GTC CGA GCC–3' (Ultramilde-Festphasensynthese bei EllaBiotech, „DMT-on“)
ODN 19	5'–GA fdC GTC CCC AAC CTC fdCGT CTG GAA GAC ACA GGC AGA TAG fdCGC TfdCG CCT C–3'

Tabelle 13: Sequenzen aller in dieser Arbeit synthetisierten DNA/RNA hybrid Oligonukleotide. Modifizierte Basen wurden Fett markiert. B = 5'–Biotin (LinkTech/ Stöckl); N = A/C/G/U.

ORN	Sequenz
ODRN 5	5'–B NNN GmNN NC–3'
ODRN 6	5'–B NNN rGNN NC–3'
ODRN 20	5'–B GG m⁶A CTG Gm⁶AC TGG m⁶ACT GG m⁶A CT–3'
ODRN 21	5'–B GGA CTG GAC TGG ACT GGA CT–3'

Tabelle 14: Sequenzen aller in dieser Arbeit synthetisierten Oloribonukleotide. Modifizierte Basen wurden Fett markiert. B = 5'–Biotin (LinkTech/ Stöckl); cIB = „Clicked“-Biotin (LinkTech/ Stöckl); N = A/C/G/U; cIU = C8-Alkyne-dU (Baseclick).

ORN	Sequenz
ORN 22	5'–GG m⁶A CU–3'
ORN 23	5'–GUG m⁶Am⁶AC UU–3'
ORN 24	5'–F UG m⁶A CCG m⁶AUG Gm⁶AG G Um⁶A G Um⁶A Um⁶AG –3'
ORN 25	5'–F UGA CCG AUG GAG GUA GUA UAG–3'
ORN 26	5'–ACU GGC UAC CUC CAU CAU AUC–3'
ORN 27	5'–cIU TTT TTT TTT TTT Tms²i⁶AC AGG–3'
ORN 28	5'–cIU TTT TTT TTT TTT TAC ms²i⁶AGG –3'
ORN 29	5'–cIU TTT TAA AAA AAA TTT Tms²i⁶AC AGG–3'
ORN 30	5'–cIU TTT TAA AAA AAA TTT TAC ms²i⁶AGG –3'
ORN 31	5'–GCT AGA TAC ms²i⁶ACC GTG ACT CU–3'
ORN 32	5'–cIB GGG GAU UGA Ams²i⁶AA UCC CC–3'
ORN 33	5'–cIB GGG GAU UGA AAA UCC CC–3'

8.3. Biochemie Methoden

8.3.1. Allgemeine Arbeitsmethoden

Puffer und Lösungen

Die in dieser Arbeit genutzten Puffer und Lösungen wurden grundsätzlich mit frischem erzeugtem Reinstwasser (Leitfähigkeit entspricht $18.2 \text{ M}\Omega \times \text{cm}$) hergestellt, im Weiteren als ddH₂O beschrieben. Chemikalien wurden von den Firmen *Merck*, *Sigma-Aldrich* und *AppliChem* in den üblichen Qualitätsstufen *puriss*, *p.a.*, *purum* und *molecular biology grade* bezogen. Zur Einstellung des pH-Wertes wurde ein *MP220* pH-Meter (*Mettler Toledo*) verwendet.

Tabelle 15: Puffer und Lösungen die in dieser Dissertation verwendet wurden.

Bezeichnung	Zusammensetzung
TBS	20 mM Tris·HCl, 150 mM NaCl, pH = 7.5
TBS-T	20 mM Tris·HCl, 150 mM NaCl, 0.1 % Tween 20, pH = 7.5
EDC-Konjugationspuffer	0.1 M MES, 0.9 M NaCl, 0.015 % NaN ₃ , pH = 4.6
PBS-Elutionspuffer	8.3 mM Natriumphosphat, 0.9 M NaCl, 0.1 M D-Sorbitol, pH = 7.2
HPA Matrix (MALDI)	44 mM di-Ammoniumhydrogencitrat, 3-Hydroxypicolinsäure (übersättigt) in MeCN/ddH ₂ O (1:1)
SDS-Laufpuffer	25 mM Tris HCl, 192 mM Glycin, 0.1% SDS

Thermoschüttler, Speed-Vac und Lyophilisierung

Die Temperatureinstellung und Durchmischung von Reaktionen in *eppendorf* Reaktionsgefäßen (2 mL und 1.5 mL) und *Falkon* Reaktionsgefäßen (15 mL) wurde durch die Thermoschüttler *Thermomixer-comfort* (*eppendorf*) erreicht.

Um Lösungsmittel zügig aus den oben genannten Reaktionsgefäßen zu entfernen wurde eine *SPD111V Speed Vac* (*Savant*) verwendet.

Wasser und Wasser/MeCN-Gemische wurden mittels flüssigem Stickstoff eingefroren und durch eine *Alpha 2-4 LD plus* Lyophille von der Firma *Christ* gefriergetrocknet.

Bidestilliertes Wasser

Reinstwasser mit der Eigenleitfähigkeit von $18.2 \text{ M}\Omega \times \text{cm}$, in dieser Arbeit als bidestilliertes Wasser (ddH₂O) beschrieben, wurde kurz vor Benutzung von einer *arium Pro DI* Anlage (*Sartorius Stedim*) frisch erzeugt.

Bradford Assay

Die Konzentrationsbestimmungen von Proteinen wurden mittels Bradford durchgeführt. Die photometrische Methode basiert auf dem Farbstoff Coomassie-Brilliant-Blau G-250. Das Absorptionsmaximum des Farbstoffes bei 465 nm verschiebt sich durch Komplexbildung mit in der

Lösung gelöster Proteine auf 595 nm. Die Komplexbildung ist nach 2 min weitgehend abgeschlossen, der Absorptionsanstieg bei 595 nm kann ausgelesen werden und die Konzentration des gelösten Proteins berechnet werden.^[242] Es wurde Bradford-Farbkonzentrat *Protein Assay Dye Reagent Concentrate (Bio-Rad)* mit ddH₂O (1:9) verdünnt. In einer Einmalküvette (*Brand*) wurde 1 µL der zu untersuchenden Proteinlösung mit 49 µL H₂O verdünnt (Konzentration idealerweise im linearen Bereich, ca. 1 µg/µL) und 950 µL der vorher verdünnten Farblösung hinzu gegeben. Nach 5 min Inkubationszeit unter Lichtausschluss wurde die Absorption bei 595 nm an einem Photometer von Eppendorf (Bio Photometer) gemessen. Um Ungenauigkeiten durch verwendete Puffer zu minimieren, wurde eine zur Probe gleiche Menge Puffer (1 µL) als *blank* eingemessen. Die Messungen wurden ohne weitere Kalibrierung des Photometers durchgeführt.

UV-Vis Spektrometrie

Die Konzentrationsbestimmung von Oligonukleotiden wurde an einem *NanoDrop ND-1000 UV/Vis-Spectrophotometer (PeqLab)* durchgeführt. Das für die Bestimmung genutzte Probenvolumen lag bei 1.0 µL und die Absorption wurde bei 260 nm gemessen.

DNA Fällung

Die DNA Probe wurde mit 10 % (v/v) Natriumacetat (3M, pH = 5.5) versetzt und für 5 min bei Raumtemperatur bei wiederholtem schütteln inkubiert. Die Lösung wurde mit 2.5 x Volumen EtOH versetzt und über Nacht bei -80 °C gefällt. Das Präzipitat wurde für 1 h bei 4 °C und 21130 g zentrifugiert. Der Überstand wurde abgenommen und das Pellet mit 70-80 % EtOH_{aq} gewaschen (2 x 1 mL; 15 min, 4 °C, 21130 g). Der Überstand wurde verworfen und das Pellet an Luft getrocknet. Das getrocknete Pellet bzw. die gefällte DNA in EtOH kann über mehrere Monate bei -20 °C gelagert werden.

8.3.2. EDC-Kopplungen

8.3.2.1. Ovalbumin

Gm-Hapten

Imject[®] OVA von *Thermo Fisher* (5.0 mg; 0.01 M Na₂HPO₄, pH=7.2) wurde in H₂O (500 µL) gelöst. Das Gm-Hapten (8.0 mg, 14.1 µmol, 1.0 Äq) wurde in DMSO (225 µL) und in EDC-Konjugationspuffer (1.0 mL) gelöst und zum *Imject*[®] OVA gegeben. Die Lösung wurde zu frischem EDC-HCl (23 mg, 120 µmol, 8.5 Äq) Pulver gegeben und für 1.5 h bei 25 °C und 300 rpm inkubiert. Von einer Zeba[™] *spin desalting Column* (5 mL) wurde der Lagerpuffer (2 min, 4 °C, 1000xg) entfernt und mit PBS-Elutions Puffer (**2x**: 2.5 mL, 2 min, 4 °C, 1000xg) äquilibriert. Die Reaktionsmischung (1.75 mL) wurde gefiltert und über die Zeba[™] Säule entsalzt (2 min, 4 °C, 1000xg). Es wurden 1.5 mL einer Protein-Lösung (1.0 µg/µL, bestimmt mittels *Bradford Assay*) erhalten.

fdC-PEG3-Hapten

Imject® OVA von Thermo Fisher (5.0 mg; 0.01 M Na₂HPO₄, pH=7.2) wurde in H₂O (500 µL) gelöst. Das fdC-PEG3-Hapten (6.0 mg, 11.4 µmol, 1.0 Äq) wurde in EDC- Konjugationspuffer (1.25 mL) gelöst und zum Imject® OVA gegeben. Die Lösung wurde zu frischem EDC·HCl (23 mg, 120 µmol, 10.5 Äq) Pulver gegeben und für 1.5 h bei 25 °C und 300 rpm inkubiert. Von einer Zeba™ *spin desalting Column* (5 mL) wurde der Lagerpuffer (2 min, 4 °C, 1000xg) entfernt und mit PBS-Elutions Puffer (**2x**: 2,5 mL, 2 min, 4 °C, 1000xg) äquilibriert. Die Reaktionsmischung (1.75 mL) wurde über die Zaba™ Säule entsalzt (2 min, 4 °C, 1000xg). Es wurden 1.5 mL einer Protein-Lösung (1.6 µg/µL, bestimmt mittels *Bradford Assay*) erhalten.

fdC-PEG1-Hapten

Imject® OVA von Thermo Fisher (4 mg; 0.01 M Na₂HPO₄, pH=7.2) wurde in H₂O (400 µL) gelöst. Das fdC-PEG1-Hapten (4.8 mg, 11.4 µmol, 1.0 Äq) wurde in EDC- Konjugationspuffer (1.0 mL) gelöst und zum Imject® OVA gegeben. Die Lösung wurde zu frischem EDC·HCl (20 mg, 104 µmol, 9.5 Äq) Pulver gegeben und für 1.5 h bei 25 °C und 400 rpm inkubiert. Von einer Zeba™ *spin desalting Column* (5 mL) wurde der Lagerpuffer (2 min, 4 °C, 1000xg) entfernt und mit PBS-Elutions Puffer (**3x**: 2,5 mL, 2 min, 4 °C, 1000xg) äquilibriert. Die Reaktionsmischung (1.4 mL) wurde über die Zaba™ Säule entsalzt (2 min, 4 °C, 1000xg). Es wurden 1.4 mL einer Protein-Lösung (2.0 µg/µL, bestimmt mittels *Bradford Assay*) erhalten.

8.3.2.2. kationisiertes BSA (cBSA)

fdC-PEG3-Hapten

Imject® cBSA von Thermo Fisher (4.0 mg; 0.1 M Na₂HPO₄, 0.15 M NaCl, pH=7.2) wurde in H₂O (400 µL) gelöst. Das fdC-PEG3-Hapten (9.0 mg, 17.1 µmol, 1.0 Äq) wurde in EDC- Konjugationspuffer (1.5 mL) gelöst und zum Imject® cBSA gegeben. Die Lösung wurde zu frischem EDC·HCl (20 mg, 104 µmol, 6.1 Äq) Pulver gegeben und für 1.5 h bei 25 °C und 400 rpm inkubiert. Von einer Zeba™ *spin desalting Column* (5 mL) wurde der Lagerpuffer (2 min, 4 °C, 1000xg) entfernt und mit PBS-Elutions Puffer (**3x**: 2,5 mL, 2 min, 4 °C, 1000xg) äquilibriert. Die Reaktionsmischung (1.9 mL) wurde über die Zaba™ Säule entsalzt (2 min, 4 °C, 1000xg). Es wurden 1.8 mL einer Protein-Lösung (2.4 µg/µL, bestimmt mittels *Bradford Assay*) erhalten.

fdC-PEG1-Hapten

Imject® cBSA von Thermo Fisher (4 mg; 0.1 M Na₂HPO₄, 0.15 M NaCl, pH=7.2) wurde in H₂O (400 µL) gelöst. Das fdC-PEG1-Hapten (6.1 mg, 11.4 µmol, 1.0 Äq) wurde in EDC- Konjugationspuffer (1.0 mL) gelöst und zum Imject® cBSA gegeben. Die Lösung wurde zu frischem EDC·HCl (20 mg, 104 µmol, 7.5 Äq) Pulver gegeben und für 1.5 h bei 25 °C und 400 rpm inkubiert. Von einer Zeba™ *spin desalting Column* (5 mL) wurde der Lagerpuffer (2 min, 4 °C, 1000xg) entfernt

und mit PBS-Elutions Puffer (**3x**: 2,5 mL, 2 min, 4 °C, 1000xg) äquiliibriert. Die Reaktionsmischung (1.4 mL) wurde über die Zaba™ Säule entsalzt (2 min, 4 °C, 1000xg). Es wurden 1.4 mL einer Protein-Lösung (2.4 µg/µL, bestimmt mittels *Bradford Assay*) erhalten.

8.3.3. SDS Polyacrylamid Gelelektrophorese (SDS-PAGE)

Um die Hapten-Protein Konjugation überprüfen zu können wurde SDS-PAGE angefertigt. Zur Größenauftrennung der Proteine wurde ein 0.75 mm dickes 10 % Trenngel gegossen, mit 100 % Isopropanol überschichtet und für 30 min ausgehärtet. Die Isopropanol Schicht wurde abgegossen und das Gel mit einem 5 % Sammelgel abgeschlossen. Die 1 µg der Protein Proben wurden mit SDS-Probenpuffer für 1 min bei 92 °C erhitzt und mit dem SDS-PA Gel aufgetrennt. Der *Color Protein Broad Standard* (#P7712S) von der Firma *NEB* wurde ohne weitere Vorbehandlung auf das Gel aufgetragen. Die Gelelektrophorese wurde mit SDS-Laufpuffer bei 150 V für 1.5 h gefahren. Zur Visualisierung wurde das Gel mit *Coomassie Brilliant Blau* gefärbt und mit einem *Amersham Imager 680* fotografiert.

Tabelle 16: SDS Gel Zusammensetzung die in dieser Arbeit genutzt wurden.

Bezeichnung	Zusammensetzung
SDS-Trenngel (10 %)	<ul style="list-style-type: none"> • Acrylamid (6.7 mL, Rotiphorese Gel 30 von <i>Carl Roth</i>) • ddH₂O (7.9 mL) • TRIS-Puffer (5 mL, 1.5 M, pH = 8.3) • SDS (200 µL, 10 % in H₂O) • APS (200 µL, 10 % in H₂O) <p>Zum Start der Polymerisierung 20 µL TEMED</p>
SDS-Sammelgel (5 %)	<ul style="list-style-type: none"> • Acrylamid (1.7 mL, Rotiphorese Gel 30 von <i>Carl Roth</i>) • ddH₂O (6.8 mL) • TRIS-Puffer (1.25 mL, 1.0 M, pH = 6.8) • SDS (100 µL, 10 % in H₂O) • APS (100 µL, 10 % in H₂O) <p>Zum Start der Polymerisierung 10 µL TEMED</p>

8.3.4. Dot Blots

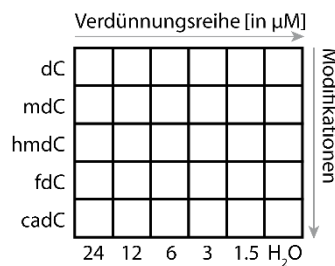


Abbildung 69: Darstellung der typischen Aufteilung der Nylon-Membran, in den durchgeführten Dot Blots zur Selektivitätsbestimmung der generierten Antikörper.

Auf eine Roti[®]-Nylon plus (Art. HP34.1) Membran (Gesamt: 2.5 cm x 3.0 cm) wurde ein Gitternetz (Kästen: 0.5 cm x 0.5 cm) aufgetragen. Auf die Membran wurden Verdünnungsreihen der zu untersuchenden Oligonukleotid Lösung (Abbildung 69 für dC, mdC, hmdC, fdC und cadC) gegeben. In der ersten Spalte wurden je 1 µL einer 24 µM Oligonukleotid Lösung und in jeder weiteren Spalte eine 1:1 Verdünnung dieser aufgetragen. In der letzten Spalte wurde zur Kontrolle H₂O aufgetragen. Die Nylon-Membran wurde bei Raumtemperatur getrocknet und entweder für 1 h bei 80 °C gebacken oder 10 min über UV Strahlung fixiert (LAS 3000; Gel-cutting“ Programm). Anschließend wurde die Membran mit 5 % Magermilchpulver in TBS-T (5 mL) für 1 h bei Raumtemperatur geblockt und mit TBS-T (3 x 5 mL) gewaschen. Die Membran wurde mit einer 1:1 Mischung des Primäre-Antikörper Überstands in TBS-T (5 mL; mit gesamt 0.1 % Magermilchpulver) über Nacht bei 4 °C inkubiert und mit TBS-T (3 x 5 mL) gewaschen. Der Sekundär-Antikörper (z.B. AntiRat HRP IgG2a) wurde 1:1000 mit TBS-T (5 mL; mit insgesamt 1 % Magermilchpulver) verdünnt und 1 h bei Raumtemperatur mit der Membran inkubiert. Anschließend wurde die Membran mit TBS-T (2 x 5 mL) und mit TBS (1 x 5 mL) gewaschen. Die Membran wurde mit einem Chemilumineszenz Meerrettichperoxidase Substrat (*SuperSignalTM West Pico PLUS Chemiluminescent Substrate* von *Thermo Scientific*) entwickelt und an einem *LAS 3000 Image Analyser* von *Fujifilm* mit Belichtungszeiten zwischen 10 s - 10 min visualisiert.

8.3.5. Übersicht getesteter Hybridomzell-Überstände

Tabelle 17: Übersicht der Hybridomzell-Überstände, welche mit Hilfe von Dot Blots auf die Selektivität gegen fdC im Vergleich zu dC und den weiteren epigenetischen Basen. Die Selektivität wurde durch eine Verdünnungsreihe eingeordnet in schwach, mittel und stark.

Bezeichnung	IgG Subklasse (Ratte)	Selektivitäten (im Dot Blot)
DNA-E R2C 5fC 2C4	IgG2c	mdC (mittel)
DNA-E R2C 5fC 7H2 + BB	IgG2c	mdC (stark), fdC (mittel), dC (schwach)
DNA-E R2C 5fC 1D3	IgG2c	mdC (stark)
DNA-E R2C 5fC 18C10	IgG2c	mdC (mittel)
DNA-E R2C 5fC 8F7 + BB	IgG2c	mdC (stark), fdC (mittel), dC (schwach)

DNA-E R2A 5fC 7H4	IgG2a	Keine Signale
DNA-E R2b 5fC 3C9	IgG2b	Keine Signale
DNA-E R2A 5fC 4H9	IgG2a	Keine Signale
DNA-E R2A 5fC 9B9	IgG2a	Keine Signale
DNA-E R2A 5fC 3F8	IgG2a	Keine Signale
DNA-E R2A 5fC 9H9	IgG2a	Keine Signale
DNA-E1 R2A 5fC 9D12	IgG2a	dC, mdC, hmdC, fdC, cadC (mittel)
DNA-E1 R2A 5fC 21G2	IgG2a	Keine Signale
DNA-E1 R2A 5fC 10F2	IgG2a	Keine Signale
DNA-E R2C 5fC 7H2	IgG2c	Keine Signale
DNA-E R2C 5fC 8F7	IgG2c	Keine Signale
DNA-E3 RG1 5fC 12B2	IgG1	fdC (stark), hmdC (mittel)
DNA-E3 RG1 5fC 9E5	IgG1	fdC (stark), hmdC (mittel)

Tabelle 18: Übersicht der Hybridomzell-Überstände, welche mit Hilfe von Dot Blots auf die Selektivität gegen Gm im Vergleich zu den 2'-O-methylierten RNA Basen (Am, Cm, Um).

Bezeichnung	IgG Subklasse (Ratte)	Selektivitäten (im Dot Blot)
DNA-F R2b mG 9C12	IgG2b	Keine Signale
DNA-F R2b mG 7A6	IgG2b	Keine Signale
DNA-F R2b mG 4A10	IgG2b	Keine Signale
DNA-F R2b mG 17H7	IgG2b	Keine Signale

8.3.6. Studien zur Oxidation eines hmdC-Oligonukleotid mit wasserlöslichen IBX-Derivaten

8.3.6.1. IBX Derivat mit negativer Ladung

Eine Lösung von p-IBX⁻ (1.3 ng-1.3 mg, 4 nmol-4 μmol, 10-10000 Äq) in Imidazol-Puffer (60 μL, 1.5 M, pH = 7-9) wurde zu einer Lösung des 13mer hmdC-Oligonukleotid (ODN **17**) (400 pmol, in 20 μL ddH₂O) gegeben. Die Reaktionsmischung wurde bei 60 °C für 0.5-45 h und 600 rpm inkubiert. Bei Probenentnahme wurden die Reaktionsmischung (80 μL) mit einer HNET₃OAc-Lösung (20 μL, 1.0 M, pH = 7) versetzt und mittels analytischer HPLC (Gradient 0 → 19 % Puffer B über 45 min; Puffer A = 0.1 M HNET₃OAc in H₂O, pH = 7; Puffer B = 0.1 M HNET₃OAc in 80 % MeCN_{aq}, pH = 7) aufgetrennt. Zur Visualisierung des hmdC- bzw. fdC-Oligonukleotid wurde die Absorption bei 260 nm gemessen. Zur Probenvorbereitung für die LC-MS/MS Messungen durch Sarah Schiffers, wurde von 35-45 min die Elution (erwartete Elution der Oligonukleotide) aufgefangen und Lyophilisiert. Die Lyophilisierte Probe wurde auf 0.57 μM verdünnt und davon 35 μL für den weiteren durch Sarah Schiffers durchgeführten Verdau benutzt.

8.3.6.2. IBX Derivate mit positiver Ladung (tertiäres Amin)

Amid-IBX⁺ und PEG3-IBX⁺

Eine wässrige Lösung des 13mer hmdC-Oligonukleotid (ODN **17**) (300 pmol, in 1.08 µL ddH₂O) wurde zu einem Imidazol-Puffer (60 µL, 1.0 M, pH = 8.5) gegeben. Zur Reaktionsmischung wurde eine Lösung von Amid-IBX⁺ (348 pg-87 ng, 600 pmol-150 nmol, 2-500 Äq, in 18,92 µL ddH₂O) bzw. PEG3-IBX⁺ (434 pg-108 ng, 600 pmol-150 nmol, 2-500 Äq in 18,92 µL ddH₂O) gegeben und bei 25-45 °C für 0.5-24 h und 600 rpm inkubiert. Bei Probenentnahme wurde die Reaktionsmischung (80 µL) mit einer HNEt₃OAc-Lösung (20 µL, 1.0 M, pH = 7) versetzt und mittels analytischer HPLC (Gradient 0 → 19 % Puffer B über 45 min; Puffer A = 0.1 M HNEt₃OAc in H₂O, pH = 7; Puffer B = 0.1 M HNEt₃OAc in 80 % MeCN_{aq}, pH = 7) aufgetrennt. Zur Visualisierung des hmdC- bzw. fdC-Oligonukleotid wurde die Absorption bei 260 nm gemessen. Zur Probenvorbereitung für die LC-MS/MS Messungen durch *Mirko Wagner*, wurde die Elution von 35-45 min (erwartete Elution der Oligonukleotide) aufgefangen und mittels EtOH gefällt. Das Pelett wurde mit ddH₂O gelöst und für den Verdau benutzt.

8.3.7. Reaktivitätsassay von SuperFly mit fdC- und fdU-Oligonukleotiden.

Die Derivatisierung der fdC und fdU Oligonukleotide wurde in einem durch Na₂EDTA (0.1 mM) und HEPES (20 mM, pH = 7.4) gepufferten System mit einem finalen Volumen von 20 µL durchgeführt. Es wurden der Katalysator *p*-Anisidin (2.0 mM) und *Superfly* (1.5 mM, von *René Rahimoff* zur Verfügung gestellt) im Puffersystem gelöst. Die Reaktion wurde mit der Zugabe von fdC-Oligonukleotid bzw. fdU-Oligonukleotiden (400 pmol, gelöst in ddH₂O) gestartet und für 5 min – 24 h bei 37 °C und 1000 rpm inkubiert. Zu den Reaktionszeitpunkten wurden 20 µL der Reaktionsmischung mit 5 µL Aceton versetzt und in flüssigem Stickstoff eingefroren. Die Proben wurden durch die Zugabe von 20 µL HNEt₃OAc-Puffer (1 M, pH = 7) aufgetaut und direkt in die analytische HPLC injiziert (Gradient über 45 min; Puffer A = 0.1 M HNEt₃OAc in H₂O, pH = 7.0; Puffer B = 0.1 M HNEt₃OAc in 80 % MeCN_{aq}, pH = 7.0; 0 → 30 % Puffer B). Die untersuchten Oligonukleotide hatten folgende Sequenzen; fdU: 5'–GTA ATG **fd**UGG TAT G–3' (von *Iacovos Michaelides* zur Verfügung gestellt) und 5'–GTA ATG **fd**CGG TAT G–3' und 5'–GTA ATG **AS**GG TAT G–3' (von *René Rahimoff* zur Verfügung gestellt). Bei der Reaktion mit abasischer Stelle wurde kein Katalysator verwendet bei einer Reaktionszeit von 40 min.^[223]

9. Abkürzungsverzeichnis

° C	Grad Celsius
3-HPA	3-Hydroxypyridin-2-carbonsäure
μ	mikro (10 ⁻⁶)
ν	Wellenzahl
A	Adenosin
Äq	Äquivalente
Å	Angström
abs.	Absolut
APS	Ammoniumperoxodisulfat
aq.	wässrig <i>aqueous</i>
ATR	abgeschwächte Totalreflexion, <i>attenuated total reflection</i>
AU	willkürliche Einheit, <i>arbitrary unit</i>
ber.	berechnet
BTT	Benzylthiotetrazol
C	Cytidin
cBSA	kationisiertes <i>bovine serum albumin</i>
CDS	Proteincodierender mRNA-Abschnitt, <i>coding sequence</i>
CED-Cl	Cyanoethoxy- <i>N,N</i> -diisopropylchlorophosphoramidit
CHES	2-(Cyclohexylamino)-ethansulfonsäure Puffer
CPG	Poröses Glasgranulat, <i>controlled pore glass</i>
CuAAC	Kupfer(I)-katalysierte Azid-Alkin-Cycloaddition
δ	chemische Verschiebung (NMR Einheit)
d	Dublett und Deuterium
Da	Dalton
DCM	Dichlormethan
ddH ₂ O	bidestilliertes Wasser (Leitfähigkeit = 18.2 MΩ x cm)
DMA	<i>N,N</i> -Dimethylannilin
DMF	<i>N,N</i> -Dimethylformamid
DMSO	Dimethylsulfoxid
DMT	Dimethoxytrityl
DNA	Desoxyribonukleinsäure, <i>desoxyribonucleic acid</i>

Abkürzungsverzeichnis

dsDNA	doppelsträngige DNA, <i>double strand DNA</i>
E.coli	<i>Escherichia coli</i>
EDC	3-(Ethyliminomethylidenamino)- <i>N,N</i> -dimethyl-propan-1-amin
EDTA	Ethylendiammintetraessigsäure
EI	Elektronenstoßionisation, <i>electron impact ionisation</i>
ELISA	<i>Enzyme-linked Immunosorbent Assay</i>
ESI	Elektrosprayionisation, <i>electrospray ionisation</i>
EtOH	Ethanol
EtOAc	Ethylacetat
<i>et. al.</i>	und andere, <i>et alii</i> bzw. <i>et aliae</i>
g	Gramm und Erdbeschleunigung
G	Guanosin
Gew.	Gewicht
GHP	<i>polypropylene membrane</i>
GlyGly	Diglycin Puffer
h	Stunde, <i>hour</i>
HEK-293	humane embryonale Nierenzellen 293, <i>human embryonic kidney 293 cells</i>
HeLa	humane Epithelzellen eines Gebärmutterhalskrebses (<i>Henrietta Lacks</i>)
HEPES	2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure Puffer
HR	hochauflösend, <i>high resolution</i>
Hz	Hertz
IBX	<i>o</i> -Iodoxybenzoesäure
IgG	Immunglobulin G
<i>i</i> Hex	<i>iso</i> -Hexan
IR	Infrarotspektroskopie
k	kilo
KO	Knockout-Stamm
L	Liter
m	milli (10^{-3}) und Multiplet
M	molar
Me	Methyl-
MeOD	d3-Methanol
MALDI	<i>matrix assisted laser desorption ionisation</i>
MeCN	Acetonitril

Abkürzungsverzeichnis

mESC	Murine embryonale Stammzelle, <i>mouse embryonic stem cell</i>
min.	Minute bzw. Minuten
MS	Massenspektrometrie
n	nano (10^{-9})
<i>n</i> BuOH	1-Butanol
NMR	Kernmagnetresonanz, <i>nuclear magnetic resonance</i>
ODN	Oligodesoxyribonukleotid
ORN	Oligoribonukleotid
OVA	Ovalbumin
p	pico (10^{-12})
p. a.	analysenrein, <i>pro analysi</i>
PAC	Phenoxyacetyl
PAGE	Polyacrylamid Gelelektrophorese
PBS	<i>phosphate buffered saline</i>
PEG	Polyethylenglycol
POCl ₃	Phosphoroxchlorid
ppm	<i>parts per million</i> (10^{-6})
q	Quartett
Q	Queuosin
RP	Umkehrphase, <i>reversed phase</i>
rt	Raumtemperatur
rpm	Umdrehungen pro Minute, <i>rounds per minute</i>
s	Sekunde und Singulett
SDS	Natriumdodecylsulfat, <i>sodium dodecyl sulfate</i>
ssDNA	einzelsträngige DNA, <i>single strand DNA</i>
t	Triplett
T	Thymin
TBS	<i>tert</i> -Butyldimethylsilyl und <i>tris-buffered saline</i>
TBS-T	<i>tris-buffered saline with Tween20</i>
TEMED	<i>N,N,N',N'</i> -Tetramethylethan-1,2-diamin
TFA	Trifluoressigsäure
THF	Tetrahydrofuran
TOF	<i>time of flight</i>
Tos	<i>p</i> -Toluolsulfonsäure

Abkürzungsverzeichnis

TRIS	<i>Tris</i> (hydroxymethyl)-aminomethan Puffer
U	Uridin
UTR	nicht translatierte mRNA-Abschnitt, <i>untranslated region</i>
UV	ultraviolett
V	Volt
WT	Wildtyp

10. Literaturverzeichnis

- [1] R. Dahm, *Hum. Genet.* **2008**, *122*, 565-581, Discovering DNA: Friedrich Miescher and the early years of nucleic acid research
- [2] P. A. Levene, *J. Biol. Chem.* **1919**, *40*, 415-424, The structure of yeast nucleic acid: IV Ammonia hydrolysis
- [3] O. T. Avery, C. M. MacLeod, M. McCarty, *J. Exp. Med.* **1944**, *79*, 137-158, Studies on the chemical nature of the substance inducing transformation of Pneumococcal types
- [4] A. D. Hershey, M. Chase, *J. Gen. Physiol.* **1952**, *36*, 39-56, Independent functions of viral protein and nucleic acid in growth of Bacteriophage
- [5] E. Chargaff, *Experientia* **1950**, *6*, 201-209, Chemical specificity of nucleic acids and mechanism of their enzymatic degradation
- [6] R. E. Franklin, R. G. Gosling, *Nature* **1953**, *171*, 740-741, Molecular configuration in sodium thymonucleate
- [7] M. H. F. Wilkins, A. R. Stokes, H. R. Wilson, *Nature* **1953**, *171*, 738-740, Molecular structure of nucleic acids: Molecular structure of deoxypentose nucleic acids
- [8] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737-738, Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid
- [9] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 964-967, Genetical implications of the structure of deoxyribonucleic acid
- [10] M. J. Bessman, I. R. Lehman, J. Adler, S. B. Zimmerman, E. S. Simms, A. Kornberg, *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 633-640, Enzymatic synthesis of deoxyribonucleic acid. III. The incorporation of pyrimidine and purine analogues into deoxyribonucleic acid
- [11] M. Meselson, F. W. Stahl, *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 671-682, The replication of DNA in *Escherichia coli*
- [12] M. F. Singer, *Science* **1968**, *162*, 433-436, 1968 Nobel laureate in medicine or physiology
- [13] F. Crick, *Nature* **1970**, *227*, 561-563, Central dogma of molecular biology
- [14] G. Cavalli, E. Heard, *Nature* **2019**, *571*, 489-499, Advances in epigenetics link genetics to the environment and disease
- [15] C. H. Waddington, C. Waddington, *J. Genet.* **1940**, *41*, 75-137, The genetic control of wing development in *Drosophila*
- [16] C. H. Waddington, *Int. J. Epidemiol.* **2011**, *41*, 10-13, The Epigenotype
- [17] R. Jaenisch, A. Bird, *Nat. Genet.* **2003**, *33*, 245, Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals
- [18] T. B. Johnson, R. D. Coghill, *J. Am. Chem. Soc.* **1925**, *47*, 2838-2844, Researches on pyrimidines. C111. The discovery of 5-methyl-cytosine in tuberculinic acid, the nucleic acid of the Tubercle Bacillus
- [19] G. R. Wyatt, *Nature* **1950**, *166*, 237-238, Occurrence of 5-methyl-cytosine in nucleic acids
- [20] R. D. Hotchkiss, *J. Biol. Chem.* **1948**, *175*, 315-332, The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography
- [21] M. Ehrlich, M. A. Gama-Sosa, L.-H. Huang, R. M. Midgett, K. C. Kuo, R. A. McCune, C. Gehrke, *Nucleic Acids Res.* **1982**, *10*, 2709-2721, Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells
- [22] M. Wagner, J. Steinbacher, T. F. J. Kraus, S. Michalakakis, B. Hackner, T. Pfaffeneder, A. Perera, M. Müller, A. Giese, H. A. Kretzschmar, T. Carell, *Angew. Chem. Int. Ed.* **2015**, *54*, 12511-12514, Age-dependent levels of 5-methyl-, 5-hydroxymethyl-, and 5-formylcytosine in human and mouse brain tissues

- [23] M. Bostick, J. K. Kim, P.-O. Estève, A. Clark, S. Pradhan, S. E. Jacobsen, *Science* **2007**, 317, 1760-1764, UHRF1 plays a role in maintaining DNA methylation in mammalian cells
- [24] H. Leonhardt, A. W. Page, H.-U. Weier, T. H. Bestor, *Cell* **1992**, 71, 865-873, A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei
- [25] M. Okano, D. W. Bell, D. A. Haber, E. Li, *Cell* **1999**, 99, 247-257, DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development
- [26] M. Okano, S. Xie, E. Li, *Nat. Genet.* **1998**, 19, 219-220, Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases
- [27] S. Pradhan, A. Bacolla, R. D. Wells, R. J. Roberts, *J. Biol. Chem.* **1999**, 274, 33002-33010, Recombinant human DNA (cytosine-5) methyltransferase: I. Expression, purification, and comparison of de novo and maintenance methylation
- [28] R. Holliday, J. Pugh, *Science* **1975**, 187, 226-232, DNA modification mechanisms and gene activity during development
- [29] A. D. Riggs, *Cytogenet. Genome Res* **1975**, 14, 9-25, X inactivation, differentiation, and DNA methylation
- [30] J. S. Griffith, H. R. Mahler, *Nature* **1969**, 223, 580-582, DNA ticketing theory of memory
- [31] F. Watt, P. L. Molloy, *Genes Dev.* **1988**, 2, 1136-1143, Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter
- [32] S. M. Iguchi-Ariga, W. Schaffner, *Genes Dev.* **1989**, 3, 612-619, CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation
- [33] M. Comb, H. M. Goodman, *Nucleic Acids Res.* **1990**, 18, 3975-3982, CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2
- [34] T. Mohandas, R. Sparkes, L. Shapiro, *Science* **1981**, 211, 393-396, Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation
- [35] C. P. Walsh, J. R. Chaillet, T. H. Bestor, *Nat. Genet.* **1998**, 20, 116-117, Transcription of IAP endogenous retroviruses is constrained by cytosine methylation
- [36] M. S. Bartolomei, A. L. Webber, M. E. Brunkow, S. M. Tilghman, *Genes Dev.* **1993**, 7, 1663-1673, Epigenetic mechanisms underlying the imprinting of the mouse H19 gene
- [37] W. Reik, A. Collick, M. L. Norris, S. C. Barton, M. A. Surani, *Nature* **1987**, 328, 248-251, Genomic imprinting determines methylation of parental alleles in transgenic mice
- [38] M. V. C. Greenberg, D. Bourc'his, *Nat. Rev. Mol. Cell Biol.* **2019**, 20, 590-607, The diverse roles of DNA methylation in mammalian development and disease
- [39] M. Monk, M. Boubelik, S. Lehnert, *Development* **1987**, 99, 371-382, Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development
- [40] J. Otani, H. Kimura, J. Sharif, T. A. Endo, Y. Mishima, T. Kawakami, H. Koseki, M. Shirakawa, I. Suetake, S. Tajima, *PLOS ONE* **2013**, 8, e82961, Cell cycle-dependent turnover of 5-hydroxymethyl cytosine in mouse embryonic stem cells
- [41] S. Zuo, R. J. Boorstein, G. W. Teebor, *Nucleic Acids Res.* **1995**, 23, 3239-3243, Oxidative damage to 5-methylcytosine in DNA
- [42] S. Kriaucionis, N. Heintz, *Science* **2009**, 324, 929-930, The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain
- [43] M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, A. Rao, *Science* **2009**, 324, 930-935, Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1
- [44] S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers, Y. Zhang, *Nature* **2010**, 466, 1129, Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification
- [45] E. Li, Y. Zhang, *Cold Spring Harb. Perspect. Biol.* **2014**, 6, DNA methylation in mammals
- [46] S. Cortellino, J. Xu, M. Sannai, R. Moore, E. Caretti, A. Cigliano, M. Le Coz, K. Devarajan, A. Wessels, D. Soprano, Lara K. Abramowitz, Marisa S. Bartolomei, F. Rambow, Maria R.

- Bassi, T. Bruno, M. Fanciulli, C. Renner, Andres J. Klein-Szanto, Y. Matsumoto, D. Kobi, I. Davidson, C. Alberti, L. Larue, A. Bellacosa, *Cell* **2011**, *146*, 67-79, Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair
- [47] T. Pfaffeneder, F. Spada, M. Wagner, C. Brandmayr, S. K. Laube, D. Eisen, M. Truss, J. Steinbacher, B. Hackner, O. Kotljarova, D. Schuermann, S. Michalakis, O. Kosmatchev, S. Schiesser, B. Steigenberger, N. Raddaoui, G. Kashiwazaki, U. Müller, C. G. Spruijt, M. Vermeulen, H. Leonhardt, P. Schär, M. Müller, T. Carell, *Nat. Chem. Biol.* **2014**, *10*, 574, Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA
- [48] T. Pfaffeneder, B. Hackner, M. Truß, M. Münzel, M. Müller, C. A. Deiml, C. Hagemeyer, T. Carell, *Angew. Chem. Int. Ed.* **2011**, *50*, 7008-7012, The discovery of 5-formylcytosine in embryonic stem cell DNA
- [49] S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, Y. Zhang, *Science* **2011**, *333*, 1300-1303, Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine
- [50] Y.-F. He, B.-Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C.-X. Song, K. Zhang, C. He, G.-L. Xu, *Science* **2011**, *333*, 1303-1307, Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA
- [51] A. Maiti, A. C. Drohat, *J. Biol. Chem.* **2011**, *286*, 35334-35338, Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites
- [52] K. Iwan, R. Rahimoff, A. Kirchner, F. Spada, A. S. Schröder, O. Kosmatchev, S. Ferizaj, J. Steinbacher, E. Parsa, M. Müller, T. Carell, *Nat. Chem. Biol.* **2017**, *14*, 72, 5-Formylcytosine to cytosine conversion by C–C bond cleavage in vivo
- [53] F. Guo, X. Li, D. Liang, T. Li, P. Zhu, H. Guo, X. Wu, L. Wen, T.-P. Gu, B. Hu, Colum P. Walsh, J. Li, F. Tang, G.-L. Xu, *Cell Stem Cell* **2014**, *15*, 447-459, Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote
- [54] M. Bachman, S. Uribe-Lewis, X. Yang, M. Williams, A. Murrell, S. Balasubramanian, *Nat. Chem.* **2014**, *6*, 1049, 5-Hydroxymethylcytosine is a predominantly stable DNA modification
- [55] M. Su, A. Kirchner, S. Stazzoni, M. Müller, M. Wagner, A. Schröder, T. Carell, *Angew. Chem.* **2016**, *128*, 11974-11978, 5-Formylcytosin ist vermutlich eine semipermanente Base an definierten Genompositionen
- [56] M. Bachman, S. Uribe-Lewis, X. Yang, H. E. Burgess, M. Iurlaro, W. Reik, A. Murrell, S. Balasubramanian, *Nat. Chem. Biol.* **2015**, *11*, 555, 5-Formylcytosine can be a stable DNA modification in mammals
- [57] Cornelia G. Spruijt, F. Gnerlich, Arne H. Smits, T. Pfaffeneder, Pascal W. T. C. Jansen, C. Bauer, M. Münzel, M. Wagner, M. Müller, F. Khan, H. C. Eberl, A. Mensinga, Arie B. Brinkman, K. Lephikov, U. Müller, J. Walter, R. Boelens, H. van Ingen, H. Leonhardt, T. Carell, M. Vermeulen, *Cell* **2013**, *152*, 1146-1159, Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives
- [58] M. Iurlaro, G. Ficiz, D. Oxley, E.-A. Raiber, M. Bachman, M. J. Booth, S. Andrews, S. Balasubramanian, W. Reik, *Genome Biol.* **2013**, *14*, R119-R119, A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation
- [59] T. T. M. Ngo, J. Yoo, Q. Dai, Q. Zhang, C. He, A. Aksimentiev, T. Ha, *Nat. Commun.* **2016**, *7*, 10813, Effects of cytosine modifications on DNA flexibility and nucleosome mechanical stability
- [60] S. Dzitoyeva, H. Chen, H. Manev, *Neurobiol. Aging* **2012**, *33*, 2881-2891, Effect of aging on 5-hydroxymethylcytosine in brain mitochondria
- [61] H. Chen, S. Dzitoyeva, H. Manev, *Restor. neurol. neuros.* **2012**, *30*, 237-245, Effect of aging on 5-hydroxymethylcytosine in the mouse hippocampus
- [62] C.-X. Song, J. Diao, A. T. Brunger, S. R. Quake, *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 4338-4343, Simultaneous single-molecule epigenetic imaging of DNA methylation and hydroxymethylation

- [63] Y. Cheng, A. Bernstein, D. Chen, P. Jin, *Exp. Neurol.* **2015**, 268, 3-9, 5-Hydroxymethylcytosine: A new player in brain disorders?
- [64] T. Prikrylova, J. Robertson, F. Ferrucci, D. Konorska, H. Aanes, A. Manaf, B. Zhang, C. B. Vågbø, A. Kuśnierczyk, K. M. Gilljam, *Sci. Rep.* **2019**, 9, 11065, 5-Hydroxymethylcytosine marks mammalian origins acting as a barrier to replication
- [65] M. W. Kellinger, C.-X. Song, J. Chong, X.-Y. Lu, C. He, D. Wang, *Nat. Struct. Mol. Biol.* **2012**, 19, 831, 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription
- [66] F. Li, Y. Zhang, J. Bai, M. M. Greenberg, Z. Xi, C. Zhou, *J. Am. Chem. Soc.* **2017**, 139, 10617-10620, 5-Formylcytosine yields DNA–protein cross-links in nucleosome core particles
- [67] S. Ji, H. Shao, Q. Han, C. L. Seiler, N. Y. Tretyakova, *Angew. Chem. Int. Ed.* **2017**, 56, 14130-14134, Reversible DNA-protein cross-linking at Epigenetic DNA marks
- [68] E.-A. Raiber, G. Portella, S. Martínez Cuesta, R. Hardisty, P. Murat, Z. Li, M. Iurlaro, W. Dean, J. Spindel, D. Beraldi, Z. Liu, M. A. Dawson, W. Reik, S. Balasubramanian, *Nat. Chem.* **2018**, 10, 1258-1266, 5-Formylcytosine organizes nucleosomes and forms Schiff base interactions with histones in mouse embryonic stem cells
- [69] A. M. Maxam, W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **1977**, 74, 560-564, A new method for sequencing DNA
- [70] F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **1977**, 74, 5463-5467, DNA sequencing with chain-terminating inhibitors
- [71] L. M. Smith, J. Z. Sanders, R. J. Kaiser, P. Hughes, C. Dodd, C. R. Connell, C. Heiner, S. B. H. Kent, L. E. Hood, *Nature* **1986**, 321, 674-679, Fluorescence detection in automated DNA sequence analysis
- [72] R. D. Mitra, J. Shendure, J. Olejnik, O. Edyta Krzymanska, G. M. Church, *Anal. Biochem.* **2003**, 320, 55-65, Fluorescent in situ sequencing on polymerase colonies
- [73] D. R. Bentley, S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton, C. G. Brown, K. P. Hall, D. J. Evers, C. L. Barnes, H. R. Bignell, J. M. Boutell, J. Bryant, R. J. Carter, R. Keira Cheetham, A. J. Cox, D. J. Ellis, M. R. Flatbush, N. A. Gormley, S. J. Humphray, L. J. Irving, M. S. Karbelashvili, S. M. Kirk, H. Li, X. Liu, K. S. Maisinger, L. J. Murray, B. Obradovic, T. Ost, M. L. Parkinson, M. R. Pratt, I. M. J. Rasoloniato, M. T. Reed, R. Rigatti, C. Rodighiero, M. T. Ross, A. Sabot, S. V. Sankar, A. Scally, G. P. Schroth, M. E. Smith, V. P. Smith, A. Spiridou, P. E. Torrance, S. S. Tzonev, E. H. Vermaas, K. Walter, X. Wu, L. Zhang, M. D. Alam, C. Anastasi, I. C. Aniebo, D. M. D. Bailey, I. R. Bancarz, S. Banerjee, S. G. Barbour, P. A. Baybayan, V. A. Benoit, K. F. Benson, C. Bevis, P. J. Black, A. Boodhun, J. S. Brennan, J. A. Bridgham, R. C. Brown, A. A. Brown, D. H. Buermann, A. A. Bundu, J. C. Burrows, N. P. Carter, N. Castillo, M. Chiara E. Catenazzi, S. Chang, R. Neil Cooley, N. R. Crake, O. O. Dada, K. D. Diakoumakos, B. Dominguez-Fernandez, D. J. Earnshaw, U. C. Egbujor, D. W. Elmore, S. S. Etchin, M. R. Ewan, M. Fedurco, L. J. Fraser, K. V. Fuentes Fajardo, W. Scott Furey, D. George, K. J. Gietzen, C. P. Goddard, G. S. Golda, P. A. Granieri, D. E. Green, D. L. Gustafson, N. F. Hansen, K. Harnish, C. D. Haudenschild, N. I. Heyer, M. M. Hims, J. T. Ho, A. M. Horgan, et al., *Nature* **2008**, 456, 53, Accurate whole human genome sequencing using reversible terminator chemistry
- [74] J. Shendure, S. Balasubramanian, G. M. Church, W. Gilbert, J. Rogers, J. A. Schloss, R. H. Waterston, *Nature* **2017**, 550, 345, DNA sequencing at 40: Past, present and future
- [75] W. W. Li, L. Gong, H. Bayley, *Angew. Chem. Int. Ed.* **2013**, 52, 4350-4355, Single-molecule detection of 5-hydroxymethylcytosine in DNA through chemical modification and nanopore analysis
- [76] B. A. Flusberg, D. R. Webster, J. H. Lee, K. J. Travers, E. C. Olivares, T. A. Clark, J. Korlach, S. W. Turner, *Nat. Methods* **2010**, 7, 461, Direct detection of DNA methylation during single-molecule, real-time sequencing
- [77] R. Shapiro, V. DeFate, M. Welcher, *J. Am. Chem. Soc.* **1974**, 96, 906-912, Deamination cytosine derivatives by bisulfite. Mechanism of the reaction
- [78] H. Hayatsu, Y. Wataya, K. Kai, S. Iida, *Biochemistry* **1970**, 9, 2858-2865, Reaction of sodium bisulfite with uracil, cytosine, and their derivatives

- [79] R. Shapiro, R. E. Servis, M. Welcher, *J. Am. Chem. Soc.* **1970**, 92, 422-424, Reactions of uracil and cytosine derivatives with sodium bisulfite
- [80] M. Frommer, L. E. McDonald, D. S. Millar, C. M. Collis, F. Watt, G. W. Grigg, P. L. Molloy, C. L. Paul, *Proc. Natl. Acad. Sci. U.S.A.* **1992**, 89, 1827-1831, A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands
- [81] R. Feil, J. Charlton, A. P. Bird, J. Walter, W. Reik, *Nucleic Acids Res.* **1994**, 22, 695-696, Methylation analysis on individual chromosomes: improved protocol for bisulphite genomic sequencing
- [82] Y. Huang, W. A. Pastor, Y. Shen, M. Tahiliani, D. R. Liu, A. Rao, *PLOS ONE* **2010**, 5, e8888, The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing
- [83] S.-G. Jin, S. Kadam, G. P. Pfeifer, *Nucleic Acids Res.* **2010**, 38, e125, Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine
- [84] M. J. Booth, E.-A. Raiber, S. Balasubramanian, *Chem. Rev.* **2015**, 115, 2240-2254, Chemical methods for decoding cytosine modifications in DNA
- [85] M. J. Booth, M. R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik, S. Balasubramanian, *Science* **2012**, 336, 934-937, Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution
- [86] M. J. Booth, T. W. B. Ost, D. Beraldi, N. M. Bell, M. R. Branco, W. Reik, S. Balasubramanian, *Nat. Protoc.* **2013**, 8, 1841, Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine
- [87] M. Yu, Gary C. Hon, Keith E. Szulwach, C.-X. Song, L. Zhang, A. Kim, X. Li, Q. Dai, Y. Shen, B. Park, J.-H. Min, P. Jin, B. Ren, C. He, *Cell* **2012**, 149, 1368-1380, Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome
- [88] M. J. Booth, G. Marsico, M. Bachman, D. Beraldi, S. Balasubramanian, *Nat. Chem.* **2014**, 6, 435, Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution
- [89] C.-X. Song, K. E. Szulwach, Q. Dai, Y. Fu, S.-Q. Mao, L. Lin, C. Street, Y. Li, M. Poidevin, H. Wu, J. Gao, P. Liu, L. Li, G.-L. Xu, P. Jin, C. He, *Cell* **2013**, 153, 678-691, Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming
- [90] L. Xu, Y.-C. Chen, J. Chong, A. Fin, L. S. McCoy, J. Xu, C. Zhang, D. Wang, *Angew. Chem. Int. Ed.* **2014**, 53, 11223-11227, Pyrene-based quantitative detection of the 5-formylcytosine loci symmetry in the CpG duplex content during TET-dependent demethylation
- [91] E.-A. Raiber, D. Beraldi, G. Ficz, H. E. Burgess, M. R. Branco, P. Murat, D. Oxley, M. J. Booth, W. Reik, S. Balasubramanian, *Genome Biol.* **2012**, 13, R69, Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase
- [92] H. Wu, X. Wu, L. Shen, Y. Zhang, *Nat. Biotechnol.* **2014**, 32, 1231, Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing
- [93] F. Neri, D. Incarnato, A. Krepelova, C. Parlato, S. Oliviero, *Nat. Protoc.* **2016**, 11, 1191, Methylation-assisted bisulfite sequencing to simultaneously map 5fC and 5caC on a genome-wide scale for DNA demethylation analysis
- [94] H. Wu, X. Wu, Y. Zhang, *Nat. Protoc.* **2016**, 11, 1081, Base-resolution profiling of active DNA demethylation using MAB-seq and caMAB-seq
- [95] K. Tanaka, A. Okamoto, *Bioorg. Med. Chem. Lett.* **2007**, 17, 1912-1915, Degradation of DNA by bisulfite treatment
- [96] N. Olova, F. Krueger, S. Andrews, D. Oxley, R. Berrens, M. Branco, W. Reik, *Genome Biol.* **2018**, 19, Comparison of whole-genome bisulfite sequencing library preparation strategies identifies sources of biases affecting DNA methylation data
- [97] M. Iurlaro, G. Ficz, D. Oxley, E.-A. Raiber, M. Bachman, M. J. Booth, S. Andrews, S. Balasubramanian, W. Reik, *Genome Biol.* **2013**, 14, R119, A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation

- [98] D. M. Woodcock, C. B. Lawler, M. E. Linsenmeyer, J. P. Doherty, W. D. Warren, *J. Biol. Chem.* **1997**, 272, 7810-7816, Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon
- [99] C. D. Laird, N. D. Pleasant, A. D. Clark, J. L. Sneed, K. M. A. Hassan, N. C. Manley, J. C. Vary, Jr., T. Morgan, R. S. Hansen, R. Stöger, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, 101, 204-209, Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules
- [100] B. Khulan, R. F. Thompson, K. Ye, M. J. Fazzari, M. Suzuki, E. Stasiek, M. E. Figueroa, J. L. Glass, Q. Chen, C. Montagna, E. Hatchwell, R. R. Selzer, T. A. Richmond, R. D. Green, A. Melnick, J. M. Greally, *Genome Res.* **2006**, 16, 1046-1055, Comparative isoschizomer profiling of cytosine methylation: The HELP assay
- [101] A. L. Brunner, D. S. Johnson, S. W. Kim, A. Valouev, T. E. Reddy, N. F. Neff, E. Anton, C. Medina, L. Nguyen, E. Chiao, C. B. Oyolu, G. P. Schroth, D. M. Absher, J. C. Baker, R. M. Myers, *Genome Res.* **2009**, 19, 1044-1056, Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver
- [102] A. Szwagierczak, A. Brachmann, C. S. Schmidt, S. Bultmann, H. Leonhardt, F. Spada, *Nucleic Acids Res.* **2011**, 39, 5149-5156, Characterization of PvuRtsII endonuclease as a tool to investigate genomic 5-hydroxymethylcytosine
- [103] M. Berney, J. F. McGouran, *Nat. Rev. Chem.* **2018**, 2, 332-348, Methods for detection of cytosine and thymine modifications in DNA
- [104] C.-X. Song, K. E. Szulwach, Y. Fu, Q. Dai, C. Yi, X. Li, Y. Li, C.-H. Chen, W. Zhang, X. Jian, J. Wang, L. Zhang, T. J. Looney, B. Zhang, L. A. Godley, L. M. Hicks, B. T. Lahn, P. Jin, C. He, *Nat. Biotechnol.* **2010**, 29, 68, Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine
- [105] D. Han, X. Lu, Alan H. Shih, J. Nie, Q. You, Meng M. Xu, Ari M. Melnick, Ross L. Levine, C. He, *Mol. Cell* **2016**, 63, 711-719, A highly sensitive and robust method for genome-wide 5hmC profiling of rare cell populations
- [106] C. Zhu, Y. Gao, H. Guo, B. Xia, J. Song, X. Wu, H. Zeng, K. Kee, F. Tang, C. Yi, *Cell Stem Cell* **2017**, 20, 720-731.e725, Single-cell 5-formylcytosine landscapes of mammalian early embryos and ESCs at single-base resolution
- [107] Y. Liu, P. Siejka-Zielińska, G. Velikova, Y. Bi, F. Yuan, M. Tomkova, C. Bai, L. Chen, B. Schuster-Böckler, C.-X. Song, *Nat. Biotechnol.* **2019**, 37, 424-429, Bisulfite-free direct detection of 5-methylcytosine and 5-hydroxymethylcytosine at base resolution
- [108] M. Weber, J. J. Davies, D. Wittig, E. J. Oakeley, M. Haase, W. L. Lam, D. Schübeler, *Nat. Genet.* **2005**, 37, 853-862, Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells
- [109] M. Ko, Y. Huang, A. M. Jankowska, U. J. Pape, M. Tahiliani, H. S. Bandukwala, J. An, E. D. Lamperti, K. P. Koh, R. Ganetzky, X. S. Liu, L. Aravind, S. Agarwal, J. P. Maciejewski, A. Rao, *Nature* **2010**, 468, 839, Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2
- [110] W. A. Pastor, U. J. Pape, Y. Huang, H. R. Henderson, R. Lister, M. Ko, E. M. McLoughlin, Y. Brudno, S. Mahapatra, P. Kapranov, M. Tahiliani, G. Q. Daley, X. S. Liu, J. R. Ecker, P. M. Milos, S. Agarwal, A. Rao, *Nature* **2011**, 473, 394, Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells
- [111] A. Lentini, C. Lagerwall, S. Vikingsson, H. K. Mjoseng, K. Douvlataniotis, H. Vogt, H. Green, R. R. Meehan, M. Benson, C. E. Nestor, *Nat. Methods* **2018**, 15, 499-504, A reassessment of DNA-immunoprecipitation-based genomic profiling
- [112] P. Boccaletto, M. A. Machnicka, E. Purta, P. Piatkowski, B. Baginski, T. K. Wirecki, V. de Crécy-Lagard, R. Ross, P. A. Limbach, A. Kotter, M. Helm, J. M. Bujnicki, *Nucleic Acids Res.* **2018**, 46, D303-D307, MODOMICS: A database of RNA modification pathways. 2017 update
- [113] W. E. Cohn, E. Volkin, *Nature* **1951**, 167, 483-484, Nucleoside-5'-phosphates from ribonucleic acid
- [114] H. M. Goodman, J. Abelson, A. Landy, S. Brenner, J. D. Smith, *Nature* **1968**, 217, 1019-1024, Amber suppression: A nucleotide change in the anticodon of a tyrosine transfer RNA

- [115] S. Kellner, J. Neumann, D. Rosenkranz, S. Lebedeva, R. F. Ketting, H. Zischler, D. Schneider, M. Helm, *ChemComm* **2014**, 50, 3516-3518, Profiling of RNA modifications by multiplexed stable isotope labelling
- [116] B.-I. Kang, K. Miyauchi, M. Matuszewski, G. S. D'Almeida, M. A. T. Rubio, J. D. Alfonzo, K. Inoue, Y. Sakaguchi, T. Suzuki, E. Sochacka, T. Suzuki, *Nucleic Acids Res.* **2017**, 45, 2124-2136, Identification of 2-methylthio cyclic N6-threonylcarbamoyladenine (ms2ct6A) as a novel RNA modification at position 37 of tRNAs
- [117] C. Dal Magro, P. Keller, A. Kotter, S. Werner, V. Duarte, V. Marchand, M. Ignarski, A. Freiwald, R.-U. Müller, C. Dieterich, Y. Motorin, F. Butter, M. Atta, M. Helm, *Angew. Chem. Int. Ed.* **2018**, 57, 7893-7897, A vastly increased chemical variety of RNA modifications containing a thioacetal structure
- [118] M. C. Weiss, F. L. Sousa, N. Mrnjavac, S. Neukirchen, M. Roettger, S. Nelson-Sathi, W. F. Martin, *Nat. Microbiol.* **2016**, 1, 16116, The physiology and habitat of the last universal common ancestor
- [119] J. Song, C. Yi, *ACS Chem. Biol.* **2017**, 12, 316-325, Chemical modifications to RNA: A new layer of gene expression regulation
- [120] F. Jühling, M. Mörl, R. K. Hartmann, M. Sprinzl, P. F. Stadler, J. Pütz, *Nucleic Acids Res.* **2009**, 37, D159-D162, tRNADB 2009: compilation of tRNA sequences and tRNA genes
- [121] S. K. Natchiar, A. G. Myasnikov, H. Kratzat, I. Hazemann, B. P. Klaholz, *Nature* **2017**, 551, 472, Visualization of chemical modifications in the human 80S ribosome structure
- [122] R. P. Perry, D. E. Kelley, *Cell* **1974**, 1, 37-42, Existence of methylated messenger RNA in mouse L cells
- [123] R. Desrosiers, K. Friderici, F. Rottman, *Proc. Natl. Acad. Sci. U.S.A.* **1974**, 71, 3971-3975, Identification of methylated nucleosides in messenger RNA from Novikoff Hepatoma cells
- [124] M. Helm, Y. Motorin, *Nat. Rev. Genet.* **2017**, 18, 275, Detecting RNA modifications in the epitranscriptome: Predict and validate
- [125] D. Dominissini, S. Moshitch-Moshkovitz, S. Schwartz, M. Salmon-Divon, L. Ungar, S. Osenberg, K. Cesarkas, J. Jacob-Hirsch, N. Amariglio, M. Kupiec, R. Sorek, G. Rechavi, *Nature* **2012**, 485, 201, Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq
- [126] Kate D. Meyer, Y. Saletore, P. Zumbo, O. Elemento, Christopher E. Mason, Samie R. Jaffrey, *Cell* **2012**, 149, 1635-1646, Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons
- [127] Y. Saletore, K. Meyer, J. Korch, I. D. Vilfan, S. Jaffrey, C. E. Mason, *Genome Biol.* **2012**, 13, 175, The birth of the Epitranscriptome: deciphering the function of RNA modifications
- [128] B. A. Elliott, H.-T. Ho, S. V. Ranganathan, S. Vangaveti, O. Ilkayeva, H. Abou Assi, A. K. Choi, P. F. Agris, C. L. Holley, *Nat. Commun.* **2019**, 10, 3401, Modification of messenger RNA by 2'-O-methylation regulates gene expression in vivo
- [129] J. A. Purslow, T. T. Nguyen, T. K. Egner, R. R. Dotas, B. Khatiwada, V. Venditti, *Biophys. J.* **2018**, 115, 1895-1905, Active site breathing of human Alkbh5 revealed by solution NMR and accelerated molecular dynamics
- [130] J. Liu, Y. Yue, D. Han, X. Wang, Y. Fu, L. Zhang, G. Jia, M. Yu, Z. Lu, X. Deng, Q. Dai, W. Chen, C. He, *Nat. Chem. Biol.* **2013**, 10, 93, A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation
- [131] X.-L. Ping, B.-F. Sun, L. Wang, W. Xiao, X. Yang, W.-J. Wang, S. Adhikari, Y. Shi, Y. Lv, Y.-S. Chen, X. Zhao, A. Li, Y. Yang, U. Dahal, X.-M. Lou, X. Liu, J. Huang, W.-P. Yuan, X.-F. Zhu, T. Cheng, Y.-L. Zhao, X. Wang, J. M. R. Danielsen, F. Liu, Y.-G. Yang, *Cell Res.* **2014**, 24, 177, Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase
- [132] S. Schwartz, Maxwell R. Mumbach, M. Jovanovic, T. Wang, K. Maciag, G. G. Bushkin, P. Mertins, D. Ter-Ovanesyan, N. Habib, D. Cacchiarelli, Neville E. Sanjana, E. Freinkman, Michael E. Pacold, R. Satija, Tarjei S. Mikkelsen, N. Hacohen, F. Zhang, Steven A. Carr, Eric S. Lander, A. Regev, *Cell Rep.* **2014**, 8, 284-296, Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites

- [133] D. P. Patil, C.-K. Chen, B. F. Pickering, A. Chow, C. Jackson, M. Guttman, S. R. Jaffrey, *Nature* **2016**, 537, 369, m6A RNA methylation promotes XIST-mediated transcriptional repression
- [134] T. Song, Y. Yang, H. Wei, X. Xie, J. Lu, Q. Zeng, J. Peng, Y. Zhou, S. Jiang, J. Peng, *Nucleic Acids Res.* **2019**, 47, 6130-6144, Zfp217 mediates m6A mRNA methylation to orchestrate transcriptional and post-transcriptional regulation to promote adipogenic differentiation
- [135] F. Aguilo, F. Zhang, A. Sancho, M. Fidalgo, S. Di Cecilia, A. Vashisht, D.-F. Lee, C.-H. Chen, M. Rengasamy, B. Andino, F. Jahouh, A. Roman, S. R. Krig, R. Wang, W. Zhang, J. A. Wohlschlegel, J. Wang, M. J. Walsh, *Cell Stem Cell* **2015**, 17, 689-704, Coordination of m6A mRNA methylation and gene transcription by ZFP217 regulates pluripotency and reprogramming
- [136] D. L. Balacco, M. Soller, *Biochemistry* **2019**, 58, 363-378, The m6A writer: Rise of a machine for growing tasks
- [137] P. Wang, Katelyn A. Doxtader, Y. Nam, *Mol. Cell* **2016**, 63, 306-317, Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases
- [138] C. Xu, X. Wang, K. Liu, I. A. Roundtree, W. Tempel, Y. Li, Z. Lu, C. He, J. Min, *Nat. Chem. Biol.* **2014**, 10, 927, Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain
- [139] P. J. Hsu, Y. Zhu, H. Ma, Y. Guo, X. Shi, Y. Liu, M. Qi, Z. Lu, H. Shi, J. Wang, Y. Cheng, G. Luo, Q. Dai, M. Liu, X. Guo, J. Sha, B. Shen, C. He, *Cell Res.* **2017**, 27, 1115, Ythdc2 is an N6-methyladenosine binding protein that regulates mammalian spermatogenesis
- [140] S. Luo, L. Tong, *Proc. Natl. Acad. Sci. U.S.A.* **2014**, 111, 13834-13839, Molecular basis for the recognition of methylated adenines in RNA by the eukaryotic YTH domain
- [141] Claudio R. Alarcón, H. Goodarzi, H. Lee, X. Liu, S. Tavazoie, Sohail F. Tavazoie, *Cell* **2015**, 162, 1299-1308, HNRNPA2B1 is a mediator of m6A-dependent nuclear RNA processing events
- [142] H. Huang, H. Weng, W. Sun, X. Qin, H. Shi, H. Wu, B. S. Zhao, A. Mesquita, C. Liu, C. L. Yuan, Y.-C. Hu, S. Hüttelmaier, J. R. Skibbe, R. Su, X. Deng, L. Dong, M. Sun, C. Li, S. Nachtergaele, Y. Wang, C. Hu, K. Ferchen, K. D. Greis, X. Jiang, M. Wei, L. Qu, J.-L. Guan, C. He, J. Yang, J. Chen, *Nat. Cell Biol.* **2018**, 20, 285-295, Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation
- [143] N. Liu, Q. Dai, G. Zheng, C. He, M. Parisien, T. Pan, *Nature* **2015**, 518, 560, N6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions
- [144] Z. Lin, P. J. Hsu, X. Xing, J. Fang, Z. Lu, Q. Zou, K.-J. Zhang, X. Zhang, Y. Zhou, T. Zhang, Y. Zhang, W. Song, G. Jia, X. Yang, C. He, M.-H. Tong, *Cell Res.* **2017**, 27, 1216, Mettl3-/Mettl14-mediated mRNA N6-methyladenosine modulates murine spermatogenesis
- [145] N. Liu, K. I. Zhou, M. Parisien, Q. Dai, L. Diatchenko, T. Pan, *Nucleic Acids Res.* **2017**, 45, 6051-6063, N6-Methyladenosine alters RNA structure to regulate binding of a low-complexity protein
- [146] W. Xiao, S. Adhikari, U. Dahal, Y.-S. Chen, Y.-J. Hao, B.-F. Sun, H.-Y. Sun, A. Li, X.-L. Ping, W.-Y. Lai, X. Wang, H.-L. Ma, C.-M. Huang, Y. Yang, N. Huang, G.-B. Jiang, H.-L. Wang, Q. Zhou, X.-J. Wang, Y.-L. Zhao, Y.-G. Yang, *Mol. Cell* **2016**, 61, 507-519, Nuclear m6A reader YTHDC1 regulates mRNA splicing
- [147] X. Wang, Z. Lu, A. Gomez, G. C. Hon, Y. Yue, D. Han, Y. Fu, M. Parisien, Q. Dai, G. Jia, B. Ren, T. Pan, C. He, *Nature* **2013**, 505, 117, N6-methyladenosine-dependent regulation of messenger RNA stability
- [148] B. M. Edens, C. Vissers, J. Su, S. Arumugam, Z. Xu, H. Shi, N. Miller, F. Rojas Ringeling, G.-I. Ming, C. He, H. Song, Y. C. Ma, *Cell Rep.* **2019**, 28, 845-854.e845, FMRP modulates neural differentiation through m6A-dependent mRNA nuclear export
- [149] I. A. Roundtree, M. E. Evans, T. Pan, C. He, *Cell* **2017**, 169, 1187-1200, Dynamic RNA modifications in gene expression regulation
- [150] E. Sendinc, D. Valle-Garcia, A. Dhall, H. Chen, T. Henriques, J. Navarrete-Perea, W. Sheng, S. P. Gygi, K. Adelman, Y. Shi, *Mol. Cell* **2019**, 75, 620-630.e629, PCIF1 catalyzes m6Am mRNA methylation to regulate gene expression

- [151] G. Jia, Y. Fu, X. Zhao, Q. Dai, G. Zheng, Y. Yang, C. Yi, T. Lindahl, T. Pan, Y.-G. Yang, C. He, *Nat. Chem. Biol.* **2011**, 7, 885, N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO
- [152] Y. Fu, G. Jia, X. Pang, R. N. Wang, X. Wang, C. J. Li, S. Smemo, Q. Dai, K. A. Bailey, M. A. Nobrega, K.-L. Han, Q. Cui, C. He, *Nat. Commun.* **2013**, 4, 1798, FTO-mediated formation of N6-hydroxymethyladenosine and N6-formyladenosine in mammalian RNA
- [153] J. Mauer, S. R. Jaffrey, *FEBS Lett.* **2018**, 592, 2012-2022, FTO, m6Am, and the hypothesis of reversible epitranscriptomic mRNA modifications
- [154] J. Mauer, X. Luo, A. Blanjoie, X. Jiao, A. V. Grozhik, D. P. Patil, B. Linder, B. F. Pickering, J.-J. Vasseur, Q. Chen, S. S. Gross, O. Elemento, F. Debart, M. Kiledjian, S. R. Jaffrey, *Nature* **2016**, 541, 371, Reversible methylation of m6Am in the 5' cap controls mRNA stability
- [155] J. Wei, F. Liu, Z. Lu, Q. Fei, Y. Ai, P. C. He, H. Shi, X. Cui, R. Su, A. Klungland, G. Jia, J. Chen, C. He, *Mol. Cell* **2018**, 71, 973-985.e975, Differential m6A, m6Am, and m1A demethylation mediated by FTO in the cell nucleus and cytoplasm
- [156] X. Zhang, L.-H. Wei, Y. Wang, Y. Xiao, J. Liu, W. Zhang, N. Yan, G. Amu, X. Tang, L. Zhang, G. Jia, *Proc. Natl. Acad. Sci. U.S.A.* **2019**, 116, 2919-2924, Structural insights into FTO's catalytic mechanism for the demethylation of multiple RNA substrates
- [157] G. Zheng, John A. Dahl, Y. Niu, P. Fedorcsak, C.-M. Huang, Charles J. Li, Cathrine B. Vågbo, Y. Shi, W.-L. Wang, S.-H. Song, Z. Lu, Ralph P. G. Bosmans, Q. Dai, Y.-J. Hao, X. Yang, W.-M. Zhao, W.-M. Tong, X.-J. Wang, F. Bogdan, K. Furu, Y. Fu, G. Jia, X. Zhao, J. Liu, Hans E. Krokan, A. Klungland, Y.-G. Yang, C. He, *Mol. Cell* **2013**, 49, 18-29, ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility
- [158] T. Yang, A. Cheong, X. Mai, S. Zou, E. C. Y. Woon, *ChemComm* **2016**, 52, 6181-6184, A methylation-switchable conformational probe for the sensitive and selective detection of RNA demethylase activity
- [159] C. Xu, K. Liu, W. Tempel, M. Demetriades, W. Aik, C. J. Schofield, J. Min, *J. Biol. Chem.* **2014**, 289, 17299-17311, Structures of human ALKBH5 demethylase reveal a unique binding mode for specific single-stranded N6-methyladenosine RNA demethylation
- [160] S. Zou, J. D. W. Toh, K. H. Q. Wong, Y.-G. Gao, W. Hong, E. C. Y. Woon, *Sci. Rep.* **2016**, 6, 25677, N6-Methyladenosine: a conformational marker that regulates the substrate specificity of human demethylases FTO and ALKBH5
- [161] T. Carell, C. Brandmayr, A. Hienzs, M. Müller, D. Pearson, V. Reiter, I. Thoma, P. Thumbs, M. Wagner, *Angew. Chem. Int. Ed.* **2012**, 51, 7110-7131, Structure and function of noncanonical nucleobases
- [162] S. Arragain, S. K. Handelman, F. Forouhar, F.-Y. Wei, K. Tomizawa, J. F. Hunt, T. Douki, M. Fontecave, E. Mulliez, M. Atta, *J. Biol. Chem.* **2010**, 285, 28425-28433, Identification of eukaryotic and prokaryotic methylthiotransferase for biosynthesis of 2-methylthio-N6-threonylcarbamoyladenosine in tRNA
- [163] V. Reiter, D. M. S. Matschkal, M. Wagner, D. Globisch, A. C. Kneutinger, M. Müller, T. Carell, *Nucleic Acids Res.* **2012**, 40, 6235-6240, The CDK5 repressor CDK5RAP1 is a methylthiotransferase acting on nuclear and mitochondrial RNA
- [164] F. Forouhar, S. Arragain, M. Atta, S. Gambarelli, J.-M. Mouesca, M. Hussain, R. Xiao, S. Kieffer-Jaquinod, J. Seetharaman, T. B. Acton, G. T. Montelione, E. Mulliez, J. F. Hunt, M. Fontecave, *Nat. Chem. Biol.* **2013**, 9, 333, Two Fe-S clusters catalyze sulfur insertion by radical-SAM methylthiotransferases
- [165] B. P. Anton, S. P. Russell, J. Vertrees, S. Kasif, E. A. Raleigh, P. A. Limbach, R. J. Roberts, *Nucleic Acids Res.* **2010**, 38, 6195-6205, Functional characterization of the YmcB and YqeV tRNA methylthiotransferases of *Bacillus subtilis*
- [166] B. S. Vold, J. M. Lazar, A. M. Gray, *J. Biol. Chem.* **1979**, 254, 7362-7367, Characterization of a deficiency of N6-(delta 2-isopentenyl)-2-methylthioadenosine in the *Escherichia coli* mutant *trpX* by use of antibodies to N6-(delta 2-isopentenyl)adenosine
- [167] K. Chen, Z. Lu, X. Wang, Y. Fu, G.-Z. Luo, N. Liu, D. Han, D. Dominissini, Q. Dai, T. Pan, C. He, *Angew. Chem. Int. Ed.* **2015**, 54, 1587-1590, High-resolution N6-methyladenosine (m6A) map using photo-crosslinking-assisted m6A sequencing

- [168] K. Hartstock, A. Rentmeister, *Chem.: Eur. J.* **2019**, *25*, 3455-3464, Mapping N6-methyladenosine (m6A) in RNA: Established methods, remaining challenges, and emerging approaches
- [169] B. F. Erlanger, S. M. Beiser, *Proc. Natl. Acad. Sci. U.S.A.* **1964**, *52*, 68-74, Antibodies specific for ribonucleosides and ribonucleotides and their reaction with DNA
- [170] N. S. Lipman, L. R. Jackson, L. J. Trudel, F. Weis-Garcia, *ILAR J.* **2005**, *46*, 258-268, Monoclonal versus polyclonal antibodies: distinguishing characteristics, applications, and information resources
- [171] P. Giehr, C. Kyriakopoulos, G. Ficiz, V. Wolf, J. Walter, *PLOS Comput. Biol.* **2016**, *12*, e1004905, The influence of hydroxylation on maintaining CpG methylation patterns: A hidden Markov model approach
- [172] P. Giehr, C. Kyriakopoulos, K. Lepikhov, S. Wallner, V. Wolf, J. Walter, *Nucleic Acids Res.* **2018**, *46*, e88, Two are better than one: HPoxBS - hairpin oxidative bisulfite sequencing
- [173] L. Shen, C.-X. Song, C. He, Y. Zhang, *Annu. Rev. Biochem.* **2014**, *83*, 585-614, Mechanism and function of oxidative reversal of DNA and RNA methylation
- [174] C.-X. Song, C. Yi, C. He, *Nat. Biotechnol.* **2012**, *30*, 1107, Mapping recently identified nucleotide variants in the genome and transcriptome
- [175] S. Fukuzawa, S. Takahashi, K. Tachibana, S. Tajima, I. Suetake, *Bioorg. Med. Chem.* **2016**, *24*, 4254-4262, Simple and accurate single base resolution analysis of 5-hydroxymethylcytosine by catalytic oxidative bisulfite sequencing using micelle incarcerated oxidants
- [176] T. Matsushita, Y. Moriyama, G. Nagae, H. Aburatani, A. Okamoto, *ChemComm* **2017**, *53*, 5756-5759, DNA-friendly Cu(ii)/TEMPO-catalyzed 5-hydroxymethylcytosine-specific oxidation
- [177] J. E. Hein, V. V. Fokin, *Chem. Soc. Rev.* **2010**, *39*, 1302-1315, Copper-catalyzed azide-alkyne cycloaddition (CuAAC) and beyond: new reactivity of copper(I) acetylides
- [178] S. S. Banerjee, N. Aher, R. Patil, J. Khandare, *J. Drug Deliv.* **2012**, *2012*, 103973-103973, Poly(ethylene glycol)-prodrug conjugates: Concept, design, and applications
- [179] M. Münzel, D. Globisch, C. Trindler, T. Carell, *Org. Lett.* **2010**, *12*, 5671-5673, Efficient synthesis of 5-hydroxymethylcytosine containing DNA
- [180] B. Steigenberger, S. Schiesser, B. Hackner, C. Brandmayr, S. K. Laube, J. Steinbacher, T. Pfaffeneder, T. Carell, *Org. Lett.* **2013**, *15*, 366-369, Synthesis of 5-hydroxymethyl-, 5-formyl-, and 5-carboxycytidine-triphosphates and their incorporation into oligonucleotides by polymerase chain reaction
- [181] Y. Fan, B. L. Gaffney, R. A. Jones, *Org. Lett.* **2004**, *6*, 2555-2557, Transient silylation of the guanosine O6 and amino groups facilitates N-acylation
- [182] K. L. Carraway, D. E. Koshland, in *Methods Enzymol.*, Vol. 25 (part B), Academic Press, **1972**, pp. 616-623.
- [183] J. C. Sheehan, J. J. Hlavka, *J. Am. Chem. Soc.* **1957**, *79*, 4528-4529, The cross-linking of gelatin using a water-soluble carbodiimide
- [184] D. G. Hoare, D. E. Koshland, *J. Am. Chem. Soc.* **1966**, *88*, 2057-2058, A procedure for the selective modification of carboxyl groups in proteins
- [185] G. T. Hermanson, in *Bioconjugate Techniques (Third Edition)* (Ed.: G. T. Hermanson), Academic Press, Boston, **2013**, pp. 127-228.
- [186] L. T. Hunt, M. O. Dayhoff, *Biochem. Biophys. Res. Commun.* **1980**, *95*, 864-871, A surprising new protein superfamily containing ovalbumin, antithrombin-III, and alpha1-proteinase inhibitor
- [187] J. A. Huntington, P. E. Stein, *J. Chromatogr. B Biomed. Sci. Appl.* **2001**, *756*, 189-198, Structure and properties of ovalbumin
- [188] E. Harlow, D. Lane, *Antibodies: a laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, **1988**.
- [189] G. AbouHaidar Mounir, G. Ivanov Ivan, *Z. Naturforsch. C* **1999**, *54*, 542, Non-enzymatic RNA hydrolysis promoted by the combined catalytic activity of buffers and magnesium ions

- [190] A. P. Nygaard, B. D. Hall, *Biochem. Biophys. Res. Commun.* **1963**, *12*, 98-104, A method for the detection of RNA-DNA complexes
- [191] K. C. Reed, D. A. Mann, *Nucleic Acids Res.* **1985**, *13*, 7207-7221, Rapid transfer of DNA from agarose gels to nylon membranes
- [192] E. W. Khandjian, *Bio/Technology* **1987**, *5*, 165-167, Optimized hybridization of DNA blotted and fixed to nitrocellulose and nylon membranes
- [193] G. M. Church, W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 1991-1995, Genomic sequencing
- [194] E. W. Khandjian, *Mol. Biol. Rep.* **1986**, *11*, 107-115, UV crosslinking of RNA to nylon membrane enhances hybridization signals
- [195] J. W. Holada, P. Magistro, *Vol. US Patent 8,461,171 B2*, US, **2013**.
- [196] R. G. Reed, *J. Biol. Chem.* **1986**, *261*, 15619-15624, Location of long chain fatty acid-binding sites of bovine serum albumin by affinity labeling
- [197] J. J. Babcock, L. Brancalion, *Int. J. Biol. Macromol.* **2013**, *53*, 42-53, Bovine serum albumin oligomers in the E- and B-forms at low protein concentration and ionic strength
- [198] F. S. Chu, H. P. Lau, T. S. Fan, G. S. Zhang, *J. Immunol. Methods* **1982**, *55*, 73-78, Ethylenediamine modified bovine serum albumin as protein carrier in the production of antibody against mycotoxins
- [199] A. Muckerheide, R. J. Apple, A. J. Pesce, J. G. Michael, *J. Immunol.* **1987**, *138*, 833, Cationization of protein antigens. I. Alteration of immunogenic properties
- [200] F. Tuorto, C. Legrand, C. Cirzi, G. Federico, R. Liebers, M. Müller, A. E. Ehrenhofer-Murray, G. Dittmar, H.-J. Gröne, F. Lyko, *EMBO J.* **2018**, *37*, e99777, Queuosine-modified tRNAs confer nutritional control of protein translation
- [201] S. Johannsson, P. Neumann, A. Wulf, L. M. Welp, H.-D. Gerber, M. Krull, U. Diederichsen, H. Urlaub, R. Ficner, *Sci. Rep.* **2018**, *8*, 8880, Structural insights into the stimulation of *S. pombe* Dnmt2 catalytic efficiency by the tRNA nucleoside queuosine
- [202] X. Wang, Z. Matuszek, Y. Huang, M. Parisien, Q. Dai, W. Clark, M. H. Schwartz, T. Pan, *RNA* **2018**, *24*, 1305-1313, Queuosine modification protects cognate tRNAs against ribonuclease cleavage
- [203] M. Müller, M. Hartmann, I. Schuster, S. Bender, K. L. Thüring, M. Helm, J. R. Katze, W. Nellen, F. Lyko, A. E. Ehrenhofer-Murray, *Nucleic Acids Res.* **2015**, *43*, 10952-10962, Dynamic modulation of Dnmt2-dependent tRNA methylation by the micronutrient queuine
- [204] H. Kasai, Z. Ohashi, F. Harada, S. Nishimura, N. J. Oppenheimer, P. F. Crain, J. G. Liehr, D. L. Von Minden, J. A. McCloskey, *Biochemistry* **1975**, *14*, 4198-4208, Structure of the modified nucleoside Q isolated from *Escherichia coli* transfer ribonucleic acid. 7-(4,5-cis-Dihydroxy-1-cyclopenten-3-ylaminomethyl)-7-deazaguanosine
- [205] X. Li, X. Xiong, K. Wang, L. Wang, X. Shu, S. Ma, C. Yi, *Nat. Chem. Biol.* **2016**, *12*, 311, Transcriptome-wide mapping reveals reversible and dynamic N¹-methyladenosine methylome
- [206] D. Dominissini, S. Nachtergaele, S. Moshitch-Moshkovitz, E. Peer, N. Kol, M. S. Ben-Haim, Q. Dai, A. Di Segni, M. Salmon-Divon, W. C. Clark, *Nature* **2016**, *530*, 441, The dynamic N¹-methyladenosine methylome in eukaryotic messenger RNA
- [207] P. J. Hsu, Q. Fei, Q. Dai, H. Shi, D. Dominissini, L. Ma, C. He, *Methods* **2019**, *156*, 85-90, Single base resolution mapping of 2'-O-methylation sites in human mRNA and in 3' terminal ends of small RNAs
- [208] A. V. Grozhik, A. O. Olarerin-George, M. Sindelar, X. Li, S. S. Gross, S. R. Jaffrey, *bioRxiv* **2019**, 648345, Antibody cross-reactivity accounts for widespread appearance of m¹A in 5' UTRs
- [209] B. E. Miner, R. J. Stöger, A. F. Burden, C. D. Laird, R. S. Hansen, *Nucleic Acids Res.* **2004**, *32*, e135, Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR
- [210] C. Waalwijk, R. A. Flavell, *Nucleic Acids Res.* **1978**, *5*, 3231-3236, MspI, an isoschizomer of hpaII which cleaves both unmethylated and methylated hpaII sites

- [211] J. W. Bodnar, W. Zempsky, D. Warder, C. Bergson, D. C. Ward, *J. Biol. Chem.* **1983**, 258, 15206-15213, Effect of nucleotide analogs on the cleavage of DNA by the restriction enzymes AluI, DdeI, HinfI, RsaI, and TaqI
- [212] G. Tamulaitis, M. Rutkauskas, M. Zaremba, S. Grazulis, G. Tamulaitiene, V. Siksnys, *Nucleic Acids Res.* **2015**, 43, 8100-8110, Functional significance of protein assemblies predicted by the crystal structure of the restriction endonuclease BsaWI
- [213] M. McClelland, M. Nelson, E. Raschke, *Nucleic Acids Res.* **1994**, 22, 3640-3659, Effect of site-specific modification on restriction endonucleases and DNA modification methyltransferases
- [214] K. Stankevičius, P. Povilionis, A. Lubys, S. Menkevičius, A. Janulaitis, *Gene* **1995**, 157, 49-53, Cloning and characterization of the unusual restriction-modification system comprising two restriction endonucleases and one methyltransferase
- [215] Y. Ohtsubo, Y. Nagata, M. Tsuda, *Sci. Rep.* **2017**, 7, 41769-41769, Efficient N-tailing of blunt DNA ends by Moloney murine leukemia virus reverse transcriptase
- [216] H. Klenow, I. Henningsen, *Proc. Natl. Acad. Sci. U.S.A.* **1970**, 65, 168-175, Selective elimination of the exonuclease activity of the deoxyribonucleic acid polymerase from Escherichia coli B by limited proteolysis
- [217] J. M. Clark, C. M. Joyce, G. P. Beardsley, *J. Mol. Biol.* **1987**, 198, 123-127, Novel blunt-end addition reactions catalyzed by DNA polymerase I of Escherichia coli
- [218] J. M. Clark, *Nucleic Acids Res.* **1988**, 16, 9677-9686, Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases
- [219] P. Upcroft, A. Healey, *Gene* **1987**, 51, 69-75, Rapid and efficient method for cloning of blunt-ended DNA fragments
- [220] M. S. Poptsova, I. A. Il'icheva, D. Y. Nechipurenko, L. A. Panchenko, M. V. Khodikov, N. Y. Oparina, R. V. Polozov, Y. D. Nechipurenko, S. L. Grokhovsky, *Sci. Rep.* **2014**, 4, 4532-4532, Non-random DNA fragmentation in next-generation sequencing
- [221] A. S. Schröder, J. Steinbacher, B. Steigenberger, F. A. Gnerlich, S. Schiesser, T. Pfaffeneder, T. Carell, *Angew. Chem. Int. Ed.* **2014**, 53, 315-318, Synthesis of a DNA promoter segment containing all four epigenetic nucleosides: 5-methyl-, 5-hydroxymethyl-, 5-formyl-, and 5-carboxy-2'-deoxycytidine
- [222] C. Platella, D. Musumeci, A. Arciello, F. Doria, M. Freccero, A. Randazzo, J. Amato, B. Pagano, D. Montesarchio, *Analytica Chimica Acta* **2018**, 1030, 133-141, Controlled pore glass-based oligonucleotide affinity support: Towards high throughput screening methods for the identification of conformation-selective G-quadruplex ligands
- [223] R. Rahimoff, O. Kosmatchev, A. Kirchner, T. Pfaffeneder, F. Spada, V. Brantl, M. Müller, T. Carell, *J. Am. Chem. Soc.* **2017**, 139, 10359-10364, 5-Formyl- and 5-carboxydeoxycytidines do not cause accumulation of harmful repair intermediates in stem cells
- [224] M. Yu, G. C. Hon, K. E. Szulwach, C.-X. Song, P. Jin, B. Ren, C. He, *Nat. Protoc.* **2012**, 7, 2159, Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine
- [225] R. E. Hardisty, F. Kawasaki, A. B. Sahakyan, S. Balasubramanian, *J. Am. Chem. Soc.* **2015**, 137, 9270-9272, Selective chemical labeling of natural T modifications in DNA
- [226] E. K. Schutsky, J. E. DeNizio, P. Hu, M. Y. Liu, C. S. Nabel, E. B. Fabyanic, Y. Hwang, F. D. Bushman, H. Wu, R. M. Kohli, *Nat. Biotechnol.* **2018**, 10.1038/nbt.4204, Nondestructive, base-resolution sequencing of 5-hydroxymethylcytosine using a DNA deaminase
- [227] P. Guo, S. Yan, J. Hu, X. Xing, C. Wang, X. Xu, X. Qiu, W. Ma, C. Lu, X. Weng, X. Zhou, *Org. Lett.* **2013**, 15, 3266-3269, Selective detection of 5-formyl-2'-deoxycytidine in DNA using a fluorogenic hydroxylamine reagent
- [228] Y.-C. Han, C. Zhang, *Tetrahedron Lett.* **2018**, 59, 3052-3064, Synthetic application of water-soluble hypervalent iodine reagents in aqueous media
- [229] A. Kommreddy, M. S. Bowsher, M. R. Gunna, K. Botha, T. K. Vinod, *Tetrahedron Lett.* **2008**, 49, 4378-4382, Expedient synthesis and solvent dependent oxidation behavior of a water-soluble IBX derivative

- [230] P. Acharya, P. Cheruku, S. Chatterjee, S. Acharya, J. Chattopadhyaya, *J. Am. Chem. Soc.* **2004**, *126*, 2862-2869, Measurement of nucleobase pKa values in model mononucleotides shows RNA–RNA duplexes to be more stable than DNA–DNA duplexes
- [231] P. J. Stevenson, A. B. Treacy, M. Nieuwenhuyzen, *J. Chem. Soc., Perkin Trans. 2* **1997**, 589-592, Preparation of Dess–Martin periodinane—the role of the morphology of 1-hydroxy-1,2-benziodoxol-3(1H)-one 1-oxide precursor
- [232] V. V. Zhdankin, A. Y. Kuposov, B. C. Netzel, N. V. Yashin, B. P. Rempel, M. J. Ferguson, R. R. Tykwinski, *Angew. Chem. Int. Ed.* **2003**, *42*, 2194-2196, IBX amides: A new family of hypervalent iodine reagents
- [233] H.-S. Jang, Y.-H. Kim, Y.-O. Kim, S.-M. Lee, J. W. Kim, W.-J. Chung, Y.-S. Lee, *Ind. Eng. Chem. Res.* **2014**, *20*, 29-36, Polymer-supported IBX amide for mild and efficient oxidation reactions
- [234] A. Y. Kuposov, D. N. Litvinov, V. V. Zhdankin, M. J. Ferguson, R. McDonald, R. R. Tykwinski, *Eur. J. Org. Chem.* **2006**, *2006*, 4791-4795, Preparation and reductive decomposition of 2-iodoxybenzenesulfonic acid. X-ray crystal structure of 1-hydroxy-1H-1,2,3-benziodoxathiole 3,3-dioxide
- [235] R. A. Moss, K. Y. Kim, S. Swarup, *J. Am. Chem. Soc.* **1986**, *108*, 788-793, Efficient catalytic cleavage of reactive phosphates by a functionalized o-iodosobenzoate surfactant
- [236] K. C. Nicolaou, P. S. Baran, Y. L. Zhong, *J. Am. Chem. Soc.* **2001**, *123*, 3183-3185, Selective oxidation at carbon adjacent to aromatic systems with IBX
- [237] T. Brückl, **2010**, 1-229, Synthesis of natural and isotope-labeled tRNA nucleosides and their mass spectrometric quantification
- [238] R. R. Edupuganti, S. Geiger, R. G. H. Lindeboom, H. Shi, P. J. Hsu, Z. Lu, S.-Y. Wang, M. P. A. Baltissen, P. W. T. C. Jansen, M. Rossa, M. Müller, H. G. Stunnenberg, C. He, T. Carell, M. Vermeulen, *Nat. Struct. Mol. Biol.* **2017**, *24*, 870, N6-Methyladenosine (m6A) recruits and repels proteins to regulate mRNA homeostasis
- [239] C. Zorbas, E. Nicolas, L. Wacheul, E. Huvelle, V. Heurgué-Hamard, D. L. J. Lafontaine, *Mol. Biol. Cell* **2015**, *26*, 2080-2095, The human 18S rRNA base methyltransferases DIMT1L and WBSCR22-TRMT112 but not rRNA modification are required for ribosome biogenesis
- [240] H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, *62*, 7512-7515, NMR chemical shifts of common laboratory solvents as trace impurities
- [241] L. C. J. Gillet, J. Alzeer, O. D. Schäfer, *Nucleic Acids Res.* **2005**, *33*, 1961-1969, Site-specific incorporation of N-(deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-AAF) into oligonucleotides using modified 'ultra-mild' DNA synthesis
- [242] M. M. Bradford, *Anal. Biochem.* **1976**, *72*, 248-254, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding